STEREOTYPE THREAT: AN EXAMINATION OF PROCESS

By

ALYSSA M. WALTERS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE
UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2000

TABLE OF CONTENTS

page

ABSTRACT iv

INTRODUCTION 1

Current Explanations for Stereotype Threat 3
Proposed Mediators of Stereotype Threat 4
The Proposed Model 6

REVIEW OF LITERATURE 12

Stereotype Threat Defined and Researched 13
Explanations for Stereotype Threat and Potential Mediators of Performance 16

THE PROPOSED MODEL 22

Self-Presentation and Interfering Thoughts 22
Distraction, Anxiety, and Social Facilitation 27
The Present Research 30

STUDY 1: OVERVIEW AND HYPOTHESES 32

Method 33
Results and Discussion 35

STUDY 2: OVERVIEW AND HYPOTHESES 43

Method 44
Results 47
Discussion 61

APPENDIX A STUDY 1 PARTICIPANT SELECTION QUESTIONNAIRE 71

APPENDIX B STUDY 1 TEST 72

APPENDIX C SPIELBERGER'S STATE ANXIETY INVENTORY 77

APPENDIX D STUDY 2 PARTICIPANT SELECTION CRITERIA....78
APPENDIX E ANTICIPATED SCORE FOR SELF & ADMINISTRATOR 79
APPENDIX F STUDY 2 DIFFICULT TEST....80
APPENDIX G STUDY 2 EASY TEST....85
APPENDIX H COGNITIVE INTERFERENCE QUESTIONNAIRE....88
APPENDIX I REVISED FEAR OF NEGATIVE EVALUATION SCALE..89
APPENDIX J SELF-CONSCIOUSNESS SCALE....90
APPENDIX K SHORT SPIELBERGER STATE ANXIETY INVENTORY 91
APPENDIX L POSITIVE AND NEGATIVE AFFECT SCALE....92
APPENDIX M OPEN-ENDED THOUGHT LISTINGS....93
APPENDIX N EXPERIMENTER TRAIT QUESTIONNAIRE....94
REFERENCES....95
BIOGRAPHICAL SKETCH....100

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

STEREOTYPE THREAT: AN EXAMINATION OF PROCESS

By

Alyssa M. Walters

December, 2000

Chair: James A. Shepperd
Major Department: Psychology

Stereotype threat is an apprehension that one will verify a negative stereotype about one's group. Such a predicament may lead targets to inadvertently confirm negative group stereotypes such as low intelligence and incompetence. This dissertation proposes a model of the processes that underlie stereotype threat and its consequences on test performance. Broadly, I investigated the extent to which poor performance among African American students is attributed to external evaluation as opposed to processes such as self-doubt. Specifically, I proposed that threatened students assume that an outgroup test administrator endorses negative stereotypes about their group, and this perception creates distraction and apprehension which ultimately hinder performance. Study 1 examined the role of apprehension due to an outgroup evaluative audience by varying the ethnicity of the test administrator. As predicted, African American students performed worse on a test than did European American students only when the test administrator was European American. Study 2 used structural equation modeling to test

the hypothesis that test administrator ethnicity influences test performance through the effects of perceptions of the administrator's expectations, distraction, and anxiety. The data support the predicted model. African American students' perceptions that a European American administrator held low performance expectations for them led to distraction and anxiety, which in turn culminated in poor test performance.

INTRODUCTION

Traditionally, research has focused on the cultivation and maintenance of prejudice. However, more recently, attention has shifted to understanding the experience of being the target of prejudice, stereotypes, and discrimination. The current study examines how the prevalence and widespread acceptance of cultural stereotypes impact an individual's daily experience. Of particular interest are people who are subjected to stereotypes that bear on intellectual ability. For example, how does the widely accepted stereotype that African Americans are unintelligent (Niemann, Jennings, Rozelle, & Baxter, 1994) impact the performance of African American students in the classroom and on important tests?

A review of the literature reveals that African Americans underperform in certain academic domains in comparison with European Americans. Specifically, African Americans consistently score lower on standardized tests than European Americans, a gap that appears before kindergarten and persists through adulthood (Jencks & Phillips, 1998). The disparity in academic performance has compelled explanations such as innate inferiority (Herrnstein & Murray, 1994) and socialization (Ogbu, 1986). Yet despite speculation, no one has found unequivocal evidence that African Americans have less innate intellectual ability than European Americans. Further, the differences in intellectual performance persist even after researchers control for socioeconomic status and opportunity (Miller, 1995). New evidence suggests that an interplay between

academic achievement and well-accepted stereotypes about African Americans can substantially hinder performance in the academic domain. This dissertation examines stereotype threat, a theory that posits a social-psychological account for the negative performance outcomes of African Americans and others who bear the stereotype of low intelligence.

Steele and Aronson (1995) define stereotype threat as "being at risk of confirming, as self-characteristic, a negative stereotype about one's group" (p. 797). The threat arises in situations in which a negative stereotype about one's group applies, and the group member fears being reduced to that stereotype. For instance, a female math student might fear being evaluated according to the stereotype that women are not good at math. If math is important to her, such a predicament can be particularly threatening and can harm her performance in class. Numerous studies have demonstrated the aversive consequences of this predicament (Aronson, Lustina, Good, Keough, Steele, & Brown, 1999; Croizet & Claire, 1998; Levy, 1996; Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995). Stereotype threatened people often inadvertently confirm the relevant stereotype.

Stereotype threat has garnered recent attention in social psychology and education and appears to produce deleterious effects in a variety of populations. The effect appears among African Americans on difficult standardized tests (Steele & Aronson, 1995), women in math (Spencer, Steele, & Quinn, 1999), the elderly on memory and motor skills tasks (Levy, 1996), and students from low socioeconomic backgrounds on intelligence tests (Croizet & Claire, 1998). Apparently, simply making a negative self-relevant stereotype salient can hinder performance within a setting that is relevant to that stereotype.

Stereotype threat is a relatively new phenomenon and the research is not extensive. While several studies have greatly enhanced our understanding of the phenomenon, important questions remain. Specifically, the causes of the performance deficit have eluded researchers. This dissertation examines why stereotype threat leads to a decline in task performance.

Current Explanations for Stereotype Threat

Explanations for stereotype threat's aversive consequences are primarily speculative and based on an assortment of assumptions. However, one can glean certain theoretical links from conclusions stated in the literature. I present these links in Figure 1.

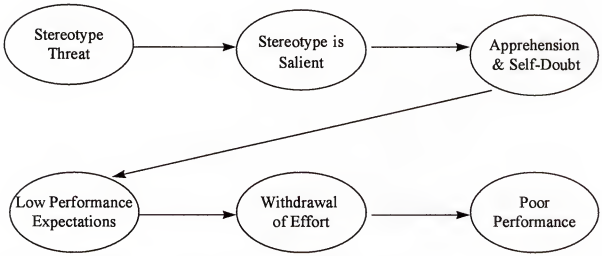

Figure 1. Suggested theoretical links for the performance decline.

In a demonstration of stereotype threat, Steele and Aronson (1995) speculated that stereotype threat may make the stereotype salient for the target which in turn leads to apprehension and self-doubt. The target becomes fearful that the stereotype may actually be true. This self-doubt leads the target to have low performance expectations which

ultimately leads to a withdrawal of effort and a subsequent decline in performance (Bandura, 1977; Pyszczynski & Greenberg, 1983). Steele and Aronson (1995) propose more broadly that stereotype threat may be due to an inefficiency of processing wherein the threatened target alternates attention between test items and the self-significance of the frustration. Several studies have examined aspects of these assumptions.

An investigation of the psychological concomitants of stereotype threat confirmed that highly subtle and implicit suggestion of the prevailing stereotype primes the stereotype for the individual (Steele & Aronson, 1995). It is also well-established that stereotype threat ultimately leads to a decline in performance (Aronson et al., 1999; Croizet & Claire, 1998; Levy, 1996; Spencer et al., 1999; Steele & Aronson, 1995). However, research on apprehension, self-doubt, and withdrawal of effort has yielded mixed evidence at best. This project examined each of these assumptions.

Proposed Mediators of Stereotype Threat

Investigations of the mediational role of anxiety in stereotype threat have yielded inconsistent results. Although several studies find heightened anxiety among stereotype threatened students (Steele & Aronson, 1995, Study 5; Spencer et al., 1999; Blascovich, Quinn, Spencer, & Steele, 1998), others find no difference in anxiety between stereotype threatened and non-stereotype threatened students (Steele & Aronson, 1995, Study 2; Aronson et al., 1999). Despite inconsistent evidence, researchers generally assume that some sort of apprehension is involved in stereotype threat, but the specific type of apprehension remains unclear. Aronson et al. (1999) have suggested that typical assessments of anxiety may be overly general and that the anxiety associated with stereotype threat may focus on evaluation apprehension.

The source of the apprehension also remains unclear. Steele and Aronson (1995) assume that apprehension represents doubt about one's capacities to perform an intellectual task. However, there is no evidence that stereotype threatened students are doubting themselves (Spencer et al., 1999; Steele & Aronson, 1995). Hence, stereotype threatened students may experience apprehension, but the apprehension unlikely involves self-doubt.

Steele and Aronson (1995) propose that a consequence of self-doubt is lowered performance expectations. Consistent with previous work, they argue that low expectations undermine effort and motivation (Bandura, 1977; Pyszczynski & Greenberg, 1983) and ultimately hinder performance. However, there is no evidence that stereotype threatened students have low expectations for themselves (Aronson et al., 1999; Steele & Aronson, 1995). Further, stereotype threatened students do not appear to withdraw effort. Indeed, some evidence suggests that stereotype threatened students may exert more effort than non-stereotype threatened students (Steele & Aronson, 1995). In sum, research finds no support for the proposition that stereotype threatened people experience self-doubt or low expectations. This evidence casts doubt on several of the theoretical links suggested by Steele and Aronson.

Considering evidence that stereotype threatened students are not withdrawing effort, Steele and Aronson (1995) postulate that stereotype threat may be moderated by a variety of factors including distraction, narrowed attention, and anxiety, which ultimately hinder the student's cognitive processing and problem-solving abilities. Additionally, Aronson et al. (1999) suggest that stereotype threat is likely motivated by a desire to confirm an image of competency. However the authors are unclear whether this motive

derives from self-doubt, or external concerns about the image of one's self and group. This dissertation addresses these theoretical possibilities and suggests a process model to account for the performance decline.

The Proposed Model

I propose that stereotype threat arises from interfering thoughts about one's ability to convey an image of competency to an important audience. Accordingly, the source of apprehension among stereotype threatened students is their perceptions of the test administrator rather than intrapsychic thoughts such as self-doubt. Surprisingly, past studies have overlooked the possible role of the test administrator in stereotype threat. Moreover, in virtually all prior studies of stereotype threat, the test administrators were not members of the stereotyped group. I propose that the academically identified student is motivated to present a competent identity image to the outgroup test administrator. However, the student assumes that the administrator endorses the stereotype of incompetence, and thus will reject this portrayal. The assumption that the student will believe that the administrator endorses the negative stereotype is consistent with the finding that people often attribute evaluative feedback to negative reactions to their stigma (Crocker, Voelkl, Testa, & Major, 1991; Crocker & Major, 1989; Kleck & Strenta, 1980; Major & Schmader, 1998).

Given that the student is motivated to present an image of competency to the outgroup test administrator, the belief that the administrator rejects this portrayal may evoke disruptive thoughts and apprehension (Schlenker, 1980; Schlenker & Leary, 1982). Presumably all students want to present an image of intellectual competence, but only students from the stereotyped group experience interfering thoughts and heightened

apprehension because only they face a test administrator who is a member of an outgroup. This assumption is consistent with Brown's (1998) finding that students of color held lower performance expectations in class when they anticipated evaluation by a European American instructor. The current model considers the experience of stigma as closely linked to the social context and the relationships with the people in that context.

As depicted in Figure 2, I suggest that the stereotype threat manipulation evokes the stereotype for participants. Students become concerned that the test administrator endorses the stereotype and thus will reject their portrayal of competency. This concern creates excessive demands on the students' attentional processes, and the students respond with a defensive narrowing of attention. Consistent with attentional models of social facilitation (Baron, 1986; Geen, 1989), whether this predicament enhances or hinders the students' performances depends on the difficulty of the task. Easy tasks require attention to a limited number of cues, while complex tasks require attention to a wide range of cues. If the task is easy, the narrowed attention facilitates performances and allows the students to convey successfully a positive identity image to the test administrator. However, if the task is difficult, the students' narrowed attention hinders performance on complex items and thwarts their positive identity portrayal. Further, with each missed item on the test, students become increasingly anxious about not successfully conveying an image of competency (Bond, 1982). The anxiety caused by an unsuccessful portrayal thus exacerbates the poor performance.

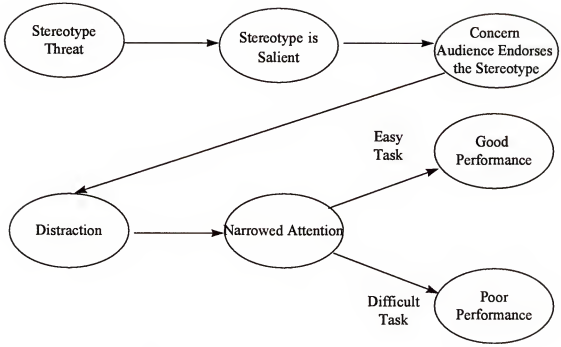

Figure 2. Proposed model for the performance decline.

Study 1: Overview and Hypotheses

The purpose of the current project is (1) to understand the impact of the test administrator on the student's performance; and more broadly (2) to test the proposed model of the processes that underlie stereotype threat and its consequences.

I conducted an exploratory study to examine the role of the test administrator on performance. Given that people who hold a stigma often assume that outgroup members endorse stereotypes about their group, I manipulated audience apprehension by varying the audience (i.e., ethnicity of the test administrator). Either an African American or a European American test administrator gave a difficult test to African American and European American students. The test administrator described the test as either

diagnostic (stereotype threatening) or nondiagnostic of intellectual ability. Consistent with previous work (Steele & Aronson, 1995), I also assessed state anxiety to examine the possible impact of test administrator ethnicity.

Hypothesis 1: When the test administrator was African American, I predicted that

(a) stereotype threatened African American students would perform equal to European American students.

(b) African American students would report anxiety equal to that of European American students.

Hypothesis 2: When the test administrator was European American, I predicted that

(a) stereotype threatened African American students would perform significantly worse on the test than European American students.

(b) African American students would report more anxiety than European American students.

(c) the ethnicity of the test administrator would not affect the performance and anxiety level of European American students.

Study 2: Overview and Hypotheses

Study 2 explored additional components of the proposed model: interfering thoughts due to perceptions of the test administrator's low expectations, and social facilitation due to the ensuing attentional deficit. As in Study 1, either an African American or European American test administrator informed students that they would take a test that was diagnostic of intellectual ability. Prior to beginning the test, students guessed the score that they believed that the test administrator believed that they would make on the test. African American and European American students completed either a

highly difficult test or a less difficult test. During the test, students completed items measuring distraction, anxiety, self-consciousness, and affect followed by open-ended thought-listings. I imbedded an easy item at the beginning and end of the difficult test to examine whether distraction increased during the test. Bond (1982) found that students become increasingly distracted on difficult tests as their performance causes them to appear incompetent to an evaluative audience.

On completion of the test, the students evaluated the study and the experimenter. Items included measures of the students' beliefs in the experimenter's endorsement of cultural stereotypes.

Hypothesis 1: When the test administrator was European American and the test was difficult, I predicted that

(a) African American students would perform worse on the test than would European American students.

(b) test administrator performance expectations, distraction, and anxiety would predict test performance as set forth in the model.

(c) African American students would be more likely than European American students to miss the easy item imbedded at the end rather than at the beginning of the test.

Hypothesis 2: When the test administrator was European American and the test was easy, I predicted that

(a) African American students would perform better on the test compared with when the test administrator was African American.

(b) test administrator performance expectations, distraction, and anxiety would predict test performance as set forth in the model.

Hypothesis 3: When the test administrator was African American, I predicted that

(a) African American students and European American students would perform equally well on the easy and difficult tests.

(b) test administrator performance expectations, distraction, and anxiety would not predict test performance.

REVIEW OF THE LITERATURE

According to Goffman (1963), stereotypes enable us to anticipate immediately a person's social identity and the attributes that correspond to that identity. However, people who are stigmatized possess a social identity that is "spoiled" or "contaminated," because they hold an attribute that differs from what is "normal" and "desirable " (p. 3). Thus, society often imputes negative attributes to stigmatized people that diverge greatly from the their actual identities. Researchers have recently explored the way targets of prejudice, stereotyping, and discrimination respond to a tainted social identity. In the past, social scientists assumed that targets of prejudice inevitably internalized society's negative views of them (Clark & Clark, 1958; Lewin, 1948; Pettigrew, 1964). However, researchers have since learned that stigmatized individuals use a variety of coping strategies to maintain self-esteem despite being socially devalued and disadvantaged (Crocker & Major, 1989; Major & Schmader, 1998). In contrast to prior theorizing, members of stigmatized groups may be active agents in experiencing success and sustaining a positive self-view.

Cultural stereotypes about stigmatized groups are prevalent and widely-known, even among people who are targets of the stereotypes and who do not endorse them (Devine, 1989; Brown, Boniecki, & Walters, 1998). Researchers have recently focused on the impact of these stereotypes on targets. For example, how might an Asian American debate team member be affected by the stereotype that Asian Americans are meek and

quiet (Niemann et al., 1994)? How does the stereotype that the elderly are poor drivers affect an elderly person during a driving test? How are African American, Latino and female students affected by stereotypes that question their intellectual capacities?Investigators have developed a program of research called stereotype threat to explore these questions.

Stereotype Threat Defined and Researched

Stereotype threat is defined as an apprehension that one will verify a negative stereotype about one's group (Steele, 1997). The threat arises in situations in which a negative stereotype about one's group applies, and the group member fears being reduced to that stereotype. For instance, a female math student might fear being evaluated according to the stereotype that women are not good at math. If math is important to her, such a predicament may be particularly threatening to her self-image. Thus, the aversive consequences of stereotype threat have garnered attention in social psychology and education research.

Researchers have demonstrated that stereotype threat often leads a target to inadvertently confirm the stereotype the individual wished to avoid. For example, Steele and Aronson (1995) told African American students that they would be taking a test that measures their intellectual ability and, thus, primed the negative ethnic stereotype of low intellect. These students performed worse on the test in comparison with European American students. However, when the researchers described the test as merely a measure of test taking style, the African American students' performances equaled the performances of European American students. Apparently, the mere introduction of a negative stereotype can hinder performance.

Researchers have demonstrated the consequences of stereotype threat with a variety of populations. College-age women performed worse than men on a difficult math test when told that the test revealed gender differences (Spencer, Steele, & Quinn, 1999). However, women who received no such information performed equal to men on the same test. Levy (1996) demonstrated that stereotype threat is not limited to stereotypes about gender and ethnicity. Elderly participants performed poorly on a memory and motor skills task only when reminded of negative stereotypes about the elderly such as dementia. Others have extended this work to social class (Croizet & Claire, 1998). The correlation between socioeconomic status and intellectual ability is often demonstrated and widely known (Neisser et al., 1996). Therefore, students from low socioeconomic backgrounds performed worse than other students on an test when it was described as diagnostic of intellectual ability. Collectively, these studies suggest that simply making a negative self-relevant stereotype salient within a setting that is relevant to that stereotype can hinder performance.

Necessary and Sufficient Conditions for Stereotype Threat

While stereotype threat can occur for a variety of populations, researchers have delimited certain necessary and sufficient conditions. First, researchers concluded that stereotype threat occurs only on extremely difficult tests (Spencer et al., 1999; Steele & Aronson, 1995). Men outperformed women on a difficult test but not on a relatively easy test (Spencer et al., 1999). However, the necessity of task difficulty remains speculative because the study did not include a stereotype threat manipulation. Second, stereotype threat occurs only for people who are highly identified with the domain (Aronson et al., 1999). Ironically, students who are most invested in academic achievement, such as

college students and college hopefuls, are most susceptible to the harmful effects of stereotype threat.

Finally, stereotype threat does not require its victims to have experienced a history of stigmatization (Aronson et al., 1999). European American male college students, a traditionally privileged cohort, performed poorly on a test when told that Asian American men tended to outperform them. This evidence suggests that anyone who belongs to a group to which a negative stereotype sometimes applies is susceptible to stereotype threat. Hence, stereotype threat focuses on situational rather than biological or sociocultural obstacles to achievement.

Psychological Concomitants of Stereotype Threat

Researchers have investigated the psychological state that accompanies the stereotype threat situation. Steele and Aronson (1995) found evidence that stereotype threat activates the stereotype for the target. African American students who received the stereotype threat condition completed more word fragments with race-related terms than did European American students and non-threatened African American students. Further, stereotype threatened students endorsed items that provided excuses for a poor test performance such as inadequate sleep and high stress. These students also indicated little interest in activities linked to the African American stereotype (for example, basketball and rap music), and only 25% revealed their race on the questionnaire. In a subsequent study, all African American students reported liking sports less than European American students, and this discrepancy increased as the African American students' negative expectations about their performance increased. These findings suggest that stereotype threatened individuals seek to dissociate themselves from a stereotypical evaluation.

Stereotype threat clearly incites a threatening predicament for targets of negative stereotypes. People may inadvertently confirm negative stereotypes about their groups despite their efforts to avoid such an evaluation. Thus far, research has provided a sense of what stereotype threatened individuals experience at the time of the threat. However, the remaining key theoretical question concerns the process by which performances falter.

Explanations for Stereotype Threat and Potential Mediators of Performance

Researchers have only speculated about the potential processes producing the performance deficits. For instance, Steele and Aronson (1995) suggest that stereotype threat primes the stereotype and leads to apprehension that the stereotype may actually be true. This self-doubt causes the student to expect to do poorly which ultimately leads to a withdrawal of effort and a subsequent decline in performance (Bandura, 1977; Pyszczynski & Greenberg, 1983).

Steele and Aronson (1995) also proposed more generally that stereotype threat may be moderated by distraction, narrowed attention, anxiety, self-consciousness, withdrawal of effort (when examined over a longer period), or excessive effort. These variables may moderate performance by limiting efficient information processing. Further, Aronson et al. (1999) suggest that stereotype threat is likely motivated by a desire to confirm an image of competency. However it is unclear whether this motive derives from self-doubt, or from concerns about the public image of one's self and group. Researchers have tested a variety of variables that may mediate the impact of stereotype threat on performance. These include self-doubt, low performance expectations, withdrawal of effort, anxiety, evaluation apprehension, distraction, and cognitive interference.

Self-Doubt and Low Performance Expectations

Steele and Aronson (1995) suggest that students may worry that the stereotype may apply to the self when they realize they are doing poorly on a difficult test. Such self-doubt ultimately leads to low performance expectations and a subsequently poor test performance. Their findings revealed that stereotype threatened students were more likely to complete word fragments with self-doubting words (LOSER, DUMB, SHAME, INFERIOR, FLUNK, HARD, & WEAK). However, it is unclear whether the students were responding to the primed stereotype of low intellect or to feelings about themselves. Stereotype threatened students also believed that their performances were worse than the performances of non-stereotype threatened students following the test (Steele & Aronson, 1995). However, these estimations reflected true performance outcomes, so it is again unclear whether this finding supports the assumption that stereotype threat leads to self-doubt. Finally, other evidence casts doubt on the contention that stereotype threat causes self-doubt. Students who received the stereotype threat reported no differences in self-efficacy (Spencer et al., 1999), confidence in responses, estimated number of correct responses (Aronson et al., 1999), academic competence or self-worth (Steele & Aronson, 1995).

Steele and Aronson speculated that self-doubt leads to low expectations which ultimately hinder performance. However, in an unpublished study reported in Steele and Aronson (1995), researchers directly manipulated high and low expectations and found no effect on performance. In sum, past research indicates that there is little reason to believe that stereotype threat produces either self-doubt or lower performance expectations.

Withdrawal of effort

Consistent with past research on self-efficacy and performance (Bandura, 1977; Pyszczynski & Greenberg, 1983), Steele and Aronson (1995) propose that self-doubt and low expectations undermine effort and motivation. However, stereotype threatened African Americans reported guessing on fewer items and took more time per item than did non-threatened students (Steele & Aronson, 1995). Other research shows that stereotype threatened students report more effort than non-threatened students (Aronson et al., 1999). Therefore, it appears that stereotype threatened students do not withdraw effort, and may in fact expend more effort than non-threatened students.

Insufficient evidence for withdrawal of effort led Steele and Aronson (1995) to speculate that inefficiency of processing may hinder performance. Specifically, the student might alternate attention between the test and the negative personal ramifications of a poor performance. Researchers have tested a number of potential sources of processing disruption.

Anxiety

Investigations of the mediational role of anxiety in stereotype threat have yielded mixed evidence. Two studies found no relationship between anxiety and stereotype threat. Aronson et al. (1999) administered the Spielberger State Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970) to European American and Asian American men after completing a difficult math test. European American men who were primed with the stereotype of Asian American superiority in math did worse than Asian American men on the test, but did not report more anxiety. Similarly, Steele and Aronson (1995) measured students' anxiety levels after completing a test described as either diagnostic or non-

diagnostic of intellectual ability. African American and European American students did not differ in reported anxiety regardless of instructions.

However, other research has found evidence for increased anxiety among stereotype threatened students. In Study 5 of Steele and Aronson (1995), stereotype threatened African American students who were asked to provide their race prior to taking the test reported more anxiety than did students whose ethnicity remained unknown. Further, Blascovich et al. (1998) found that the stereotype threat manipulation predicted increases in direct, physiological measures of anxiety such as blood pressure. Spencer et al. (1999) administered Spielberger's State Anxiety Scale after the stereotype threatening instructions and prior to the test. They found that the stereotype threat instructions increased anxiety and that anxiety predicted performance. However, anxiety only partially mediated the effect of the stereotype threat instructions on performance.

In sum, measures of anxiety have revealed group differences primarily when administered prior to the test which suggests that the stereotype threat instructions may produce some immediate arousal. However, this initial arousal did not appear to be sufficient to cause the performance decline and may actually subside on completion of the test. Further, stereotype threat may induce a more specific form of arousal than that which is detected by measures used thus far.

Other research has examined evaluation apprehension as a possible mediator of the effect of stereotype threat on performance. Studies have found limited support for evaluation apprehension. Spencer et al. (1999) assessed evaluation apprehension after the stereotype threat manipulation and prior to administering the test and found no differences on the scale among stereotype threatened and non-threatened students. Evaluation

apprehension predicted test performance, but did not account for the effect of stereotype threat on performance. Other researchers found no support for evaluation apprehension as a mediator of stereotype threat on performance when administered after the test (Aronson et al., 1999). However, in response to a single questionnaire item, stereotype threatened students expressed more concern about what the experimenter was thinking of them than non-threatened students. Aronson et al. (1999) concluded that evaluation apprehension is likely the key to stereotype threat rather than generalized anxiety and perhaps traditional measures of state anxiety are overly general. Although investigations of the mediational roles of state anxiety and evaluation apprehension have yielded mixed evidence, researchers generally agree that some sort of apprehension is involved in stereotype threat.

Distraction

Stereotype threat may involve interference with the performer's attentional processes. Specifically, distraction may prevent a student from attending to and utilizing information necessary to perform a difficult test. Research supports the notion that apprehension causes distraction during tests (Wine, 1971; Sarason, 1972). Apprehensive students become preoccupied with task-irrelevant thoughts and consequently cannot engage in the cognitive processing necessary to perform well on tests. Further, "pressure variables" such as ego-involvement and audience apprehension induce performance decrements on cognitive tasks by diverting attention away from task-relevant information (Baumeister & Showers, 1986; Bond, 1982). Thus, on complex tasks, stereotype threatened students may attend to factors related to the evaluative meaning of the task performance rather than the content of the task.

Steele and Aronson (1995) proposed that distraction, narrowed attention, self-consciousness or cognitive interference might mediate the impact of stereotype threat on performance by hindering efficient processing. The researchers tested only cognitive interference by administering the Cognitive Interference Questionnaire (Sarason, 1980) after stereotype threatened and non-threatened students completed an test. They found no reported differences in cognitive interference between stereotype threatened and non-threatened students. However, they assessed cognitive interference at the end of the study. With evaluative pressure and task demands no longer present, distracting thoughts may have diminished. Thus, past work has failed to support the mediational role of attentional processes in stereotype threat, but I propose that they remain a plausible factor in the performance decrement.

In sum, several of the theoretical links posited by Steele and Aronson (1995) have not received support. Stereotype threat does not appear to involve self-doubt, low expectations or withdrawal of effort. However, I believe that their broader speculation that stereotype threat involves inefficient processing due to mechanisms such as distraction and apprehension merits a more comprehensive examination. This dissertation reexamines factors that hinder the stereotype threatened student's ability to process information in a test taking situation.

THE PROPOSED MODEL

I propose that stereotype threat arises from interfering thoughts and apprehension about one's ability to convey an image of competency to an important audience. Specifically, I believe that the academically identified student is motivated to present a competent identity image to the outgroup test administrator. However, the student fears that the administrator endorses the stereotype of incompetence and thus will reject this portrayal.

Self-Presentation and Interfering Thoughts

People are motivated to present a favorable impression to an audience (Schlenker, 1980). The audience may be external, imagined or one's self. Aronson et al. (1999) asserted that stereotype threat may involve a self-presentational motive to sustain a self-image of competency. Such a motive would not be unique to members of stigmatized groups, but would apply to any person who is invested in appearing successful in a given domain. The researchers found support for this proposition by demonstrating that European American males were also susceptible to stereotype threat. I also maintain that stereotype threat involves self-presentational motives. However, I suggest that stereotype threat involves a desire to convince an external audience rather than one's self of one's competence.

Given that students who are susceptible to stereotype threat report academic endeavors to be ego-relevant (Aronson et al., 1999), these students may actively regulate

an image of competency to the test administrator. Steele and Aronson (1995) demonstrated that stereotype threatened students actively constructed and maintained an identity to the experimenter. For instance, stereotype threatened students endorsed itemsthat provided excuses for a poor performance on an impending test such as inadequate sleep and high stress. These students also reported little interest in activities associated with the African American stereotype such as basketball and rap music. Thus, stereotype threatened students appear to convey an identity image to the test administrator that will impede a stereotypical evaluation.

However, when people are motivated to present an image, but doubt that the audience will believe the impression, they may become apprehensive about their performance (Schlenker, 1980; Schlenker & Leary, 1982). Within the student/instructor interaction, disruptive thoughts may arise if students feel that the instructor doubts their ability to display competence. I propose that stereotype threatened students believe that the test administrator will reject the portrayal of competence, because the student assumes that the test administrator endorses the stereotype of incompetence. In some interactions, test administrators will explicitly state their stereotypical beliefs. However, more often, students infer the test administrator's stereotypical views. This notion is consistent with evidence that people often attribute evaluative feedback to negative reactions to their stigma (Crocker, Voelkl, Testa, & Major, 1991; Crocker & Major, 1989; Kleck & Strenta, 1980; Major & Schmader, 1998).

Stigma and Attributions for Feedback

Several studies reveal that people with a stigma use a variety of strategies to cope with negative appraisals from outgroup members (Crocker & Major, 1989; Major &

Schmader, 1998). Crocker and Major (1989) offer three strategies used to combat the adverse effects of prejudice and discrimination: Members of stigmatized groups (1) attribute negative outcomes to prejudice based on the stigma; (2) devalue domains in which their group fares poorly; and (3) make ingroup comparisons for outcomes. These strategies allow stigmatized people to maintain self-esteem despite negative evaluations.

Researchers have demonstrated the tendency for stigmatized people to attribute negative feedback to prejudice. Crocker et al. (1991) found that African American students were more likely to attribute positive and negative feedback to prejudice when they believed their group identity was known to the evaluator than when they believed their identity was unknown. Having their identity known buffered the self-esteem of African American students from negative feedback. Similarly, Kleck and Strenta (1980) led participants to erroneously believe that they had a stigmatizing facial disfigurement. These participants interpreted all behaviors toward them as negative reactions to their stigma.

These studies collectively suggest that students who possess a stigma are apt to assume that others are evaluating them according to their stigma. They further suggest that, in academic settings, we can expect students to believe that an outgroup instructor will evaluate them according to negative group stereotypes. For example, Brown (1998) found that students of color held lower performance expectations in class when they anticipated evaluation by a European American instructor. In contrast, students of color who believed that they would be evaluated by an instructor who shared their ethnicity anticipated a high performance in the class and a fair evaluation by the instructor.

Thus, given that the feedback from an outgroup evaluator may be seen as biased by prejudice, a test administrator from an outgroup can be a source of distraction. Surprisingly, past research has not addressed the student's experience in the interaction with the test administrator in the experimental setting (Aronson et al., 1999; Croizet & Claire, 1998; Levy, 1996; Spencer et al., 1999; Steele & Aronson, 1995). Typically, the test administrator has not shared the group membership of the participant.

Test Administrators in Past Studies

In virtually all prior studies of stereotype threat, the test administrators were not members of the stereotyped group. In Steele and Aronson's (1995) five studies, a European American experimenter administered the test and stereotype threat instructions to African American students. In Spencer et al.'s (1999) first two studies on women and stereotype threat, the experimenter was male. In Levy's (1996) work on stereotype threat and the elderly, the test administrator was a college-aged student. In sum, the test administrators in previous studies of academic performance represented the outgroup for the stereotype-threatened student, and consequently, implicitly endorsed the negative group stereotype.

While the test administrators did not share the group membership of the stereotype threatened group in a majority of studies, two exceptions exist. In Aronson et al. 's (1999) study 1, a European American female primed European American males with the stereotype of Asian American superiority in math. Also, in Spencer et al.'s (1999) study 3, a female test administrator gave the instructions and test to male and female students. However, I propose that audience induced distraction arises from the assumption that the test administrator endorses the stereotype, a behavior that usually albeit not exclusively

pertains to outgroup members. In both studies, the ingroup test administrator implicitly endorsed the stereotype by describing group differences.

Further, European American female test administrators may represent an exceptional case to this scenario. Given that an overwhelming number of instructors are European American women, they may represent prototypical instructors and may not be immune to holding gender-biases. In sum, when administrators in past studies have not shared the students' stereotypes, the students may have assumed that they were being evaluated according to the stereotype. In the two cases in which the administrator shared the students' group membership, the administrator provided instruction that implied that they endorsed the stereotype.

In this dissertation, I will examine stereotype threat among African American students. I propose that when the test administrator is European American, the student will assume that the test administrator endorses the stereotype that African Americans are intellectually inferior. Presumably both European American and African American students will want to present an image of intellectual competence, but only the African American students will experience interfering thoughts when the administrator is European American. Thus, I believe that when the test administrator is African American, African American students will not experience increased distraction. Steele and Aronson (1995) maintain that accommodating students with the administrator will unlikely be enough to override the threat of the test taking setting. However, they have not tested this assumption. I suggest that an outgroup administrator plays a pivotal role in the process of stereotype threat by inadvertently rejecting an important impression that the student hopes to convey.

I suggest that interfering thoughts during the test will create a condition of stimulus overload for the student. The student will respond with a defensive narrowing of attention. Consistent with attentional models of social facilitation (Baron, 1986; Geen, 1989), whether this predicament enhances or hinders the student's performance depends on the difficulty of the task. I predict that the performance outcome of stereotype threat will mimic that of traditional models of social facilitation.

Distraction, Anxiety, and Social Facilitation

Theories of social facilitation address the facilitating and inhibitory influence of audiences on behavior. From its inception until the late seventies, studies of social facilitation derived almost exclusively from the drive theoretical approach. Drive theories maintain that arousal is key to performance outcomes. Zajonc's dominant response model (1965) asserted that the physical presence of others increases arousal. This arousal functions as a drive that facilitates performance on easy or well-learned tasks and hinders performance on difficult or novel tasks. Cottrell (1968) recognized the impact of evaluative concerns on drive and argued that audiences and coactors elicit arousal only when they perform an evaluative role. These traditional drive models assume that an evaluative audience ignites drive to perform the dominant response. However, critics maintain that drive models are essentially descriptive and fail to explain the actual processes which link the arousal to the participant's performance (Baumeister & Showers, 1986; Geen, 1989).

A second class of theories are based on the performer's attentional processes. These alternative theories have received considerable support and have proffered process-oriented explanations for performance outcomes (Baumeister & Showers, 1986). Most

notably, Baron (1986) has suggested a theoretical approach to social facilitation based on contemporary studies of information processing and attention. According to his Distraction Conflict Theory (Baron, 1986), cognitive overload mediates the effect of an evaluative audience on performance. Specifically, overtaxed attention requires an individual to create a system of attentional priorities. Whether this narrowing of attention facilitates or hinders performance depends on the difficulty of the task. Easy tasks require attention to a limited number of cues, while complex tasks require attention to a wide range of cues. Hence, if the task is easy, the narrowed attention will facilitate performance and allow the student to convey successfully a positive identity image to the test administrator. However, if the task is difficult, the student's narrowed attention will hinder performance on complex items and thwart the positive identity portrayal.

I propose that all students will experience pressure to perform well in the presence of the audience. However, stereotype threatened students performing in the presence of an outgroup administrator confront the added task of concern over the administrator's low expectations for them. Thus, students who experience this additional distraction will experience more severe cognitive overload than students who share the test administrator's group membership and thus, do not experience stereotype threat. I believe that students respond to the increased cognitive load by narrowing their attention to a select number of cues. This narrowing of attention facilitates performance on easy tasks by eliminating distracters, and hinders performance on difficult tasks by diminishing attention to necessary cues.

Finally, Bond (1982) offers a self-presentational account of social facilitation. He identifies inferred evaluation and subsequent attentional processes as the determinants of

performance outcomes. Specifically, a performer actively regulates an image to the audience. If this image is not successfully conveyed, the performer becomes anxious, or according to Goffman (1959), loses face. Anxiety further diverts the student's attention away from the task. This anxiety becomes self-exacerbating by making the performance progressively more unfavorable. The performer engages in a moment-by-moment self-evaluation to monitor the success or failure of the performance and the implicit evaluation. Therefore, on a difficult test, Bond (1982) speculates that a student will become anxious and progressively distracted. Each missed item implies negative feedback and a failed presentation of competency. Bond (1982) demonstrated that easy items imbedded at the end of a difficult test were likely to be missed when a student's attentional capacities were overtaxed by concern over evaluations from an audience.

I propose that the stereotype threatened students experience excessive interfering thoughts, and this distraction becomes increasingly worse as the test progresses. With each missed item on the test, the student becomes increasingly anxious about not successfully conveying an image of competency and the distraction will intensify (Bond, 1982). I predicted that stereotype threatened students would miss more items at the end of an test than at the beginning of an test.

Task difficulty

Stereotype threat appears to affect students only on difficult tasks (Spencer et al., 1999). Men outperformed women on a difficult math test, but men and women performed equally well on an easy math test. However, the researchers did not manipulate stereotype threat, so the interplay between task difficulty and stereotype threat remains unclear. Steele and Aronson (1995) maintain that only difficult tasks will be sufficiently frustrating

to lead to self-doubt. I also propose that only difficult tasks will produce a decline in performance, however the decline will stem from an attentional overload and social embarrassment rather than self-doubt. No investigation has tested task difficulty as a moderator. In Study 1, I manipulated task difficulty to investigate the possible role of attention and social facilitation as a cause for performance effects due to stereotype threat.

The Present Research

I propose that describing a test as diagnostic of intellectual ability causes the stereotype of low intelligence to become salient for African American students (see Figure 3). These academically-identified students hope to convey images of intellectual competency to the test administrator. When the administrator is European American, the stereotype threatened African American students assume that the administrator will adhere to the stereotype of low intelligence and reject their portrayal of competence. This assumption leads to excessive cognitive load.

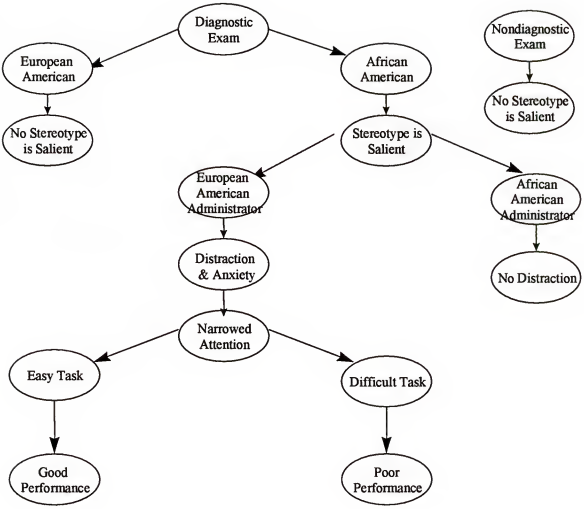

Figure 3. Proposed model for the performance decline.

The perceptions of the test administrator lead to interfering cognitions and force the students to narrow their attention to select information on the test. On a difficult test, the students are unable to attend to cues essential for a successful performance and performance suffers. On an easy test, narrowed attention facilitates performance by eliminating distracters. Finally, on a difficult task, the students become anxious about their performance and the evaluative meaning of that performance in lieu of the task.

STUDY 1: OVERVIEW AND HYPOTHESES

Study 1 examined the premise that the test administrator is a source of apprehension and responsible for the decrement in performance among stereotype threatened students. Study 1 manipulated the ethnicity (European American vs. African American) of the test administrator. Presumably, people who hold a stigma will assume that outgroup members (e.g., European Americans) endorse stereotypes about their group (Crocker, Voelkl, Testa, & Major, 1991; Crocker & Major, 1989; Kleck & Strenta, 1980; Major & Schmader, 1998). Given that some past research has found evidence that anxiety plays a role (Aronson et al., 1999; Spencer at el., 1999; Steele & Aronson, 1995), students completed a measure of state anxiety to explore its relation to test administrator ethnicity.

When the test administrator was African American, I predicted that stereotype threatened African American students' test performances and anxiety levels would parallel that of European American students. However, when the test administrator was European American, I predicted that stereotype threatened African American students would perform significantly worse on the test than European American students and would report more anxiety than European American students. I expected that the ethnicity of the test administrator would not affect the performance and anxiety level of European American students.

Method

Participants

Participants were 66 European American and 49 African American Introductory Psychology students collected across three consecutive semesters. Students received course credit for participating. Students were selected whose interest in verbal-oriented course work fell in the top tertile of students (see Appendix A).

The design was a 2 (Student ethnicity: European American vs. African American) X 2 (Test administrator ethnicity: European American vs. African American) X 2 (Test Description: Diagnostic vs. Nondiagnostic) between-subjects factorial.

Procedure

The test administrator scheduled up to four participants to arrive at the lab per session. Participants seated in separate cubicles completed a short information sheet that asked about year in school, major, and enjoyment of verbal-oriented courses (see Appendix A). Only data from students who reported interest in verbal-oriented course work were used for the study. Data from students who fell in the bottom and middle tertile on the interest item were excluded from the study. During the second semester of data collection, the experimenter recruited participants during phone interviews and only invited students to participate who reported the top tertile level of interest in verbal-oriented course work ($n = 25$).

A European American or African American female test administrator informed participants that they would take a short, very difficult test. Following Steele and Aronson's (1995) script and methodology, the administrator described the test as either

diagnostic or non-diagnostic of intellectual ability. Prior to taking the test, a subset of participants (n=80) recorded the score that they anticipated making on the test. Several participants completed the study before this measure was incorporated into the study.

Participants completed a 17-item test extracted from the most difficult items of the verbal section of the Graduate Record Examination (see Appendix B). The items were identical to those used by Steele and Aronson (1995). On completion of the test, participants completed the state form of the Spielberger's Anxiety Inventory (Spielberger et al, 1970; see Appendix C). The 20-item scale assesses respondents' state anxiety. Students responded to each item using a 4-point scale from 1 (Not at all) to 4 (Very much so) with higher responses reflecting greater state anxiety (Cronbach's $\alpha = .56$).

In prior studies of stereotype threat, Spielberger's State Anxiety Inventory yielded both supportive (Spencer at el., 1999; Steele & Aronson, 1995) and non-supportive evidence (Aronson et al.,1999; Steele &Aronson, 1995) of the mediational role of anxiety in performance. Aronson et al. (1999) suggest that the scale may be too general to tap specific processes involved in stereotype threat. I used the scale with the intention of determining the items within the scale that address positive and negative statements about affect and anxiety, in hopes of reconciling the inconsistent findings from previous reports.

After completing the study, a subset of students responded to the open-ended question, "What were the experimenters measuring?" and indicated whether the test was diagnostic or non-diagnostic of their intellectual ability. Finally, students were fully debriefed as to the purpose of the study and were able to request their actual scores on the test.

Results and Discussion

Manipulation Checks

A subset of African American participants ($n = 20$) received two manipulation check questions. Students provided an open-ended description of the primary purpose of the study, and responded to a dichotomous question asking whether the test administrator was or was not interested in knowing their intellectual ability. Manipulation checks revealed that students believed that the test was measuring intellectual ability regardless of the administrator's instructions, $\chi 2\ (1, N = 20) < 1, p = .37$. When the administrator described the test as nondiagnostic of ability, 73% of the students believed that the test was diagnostic of ability. Ninety percent of the students who received the diagnostic instructions believed the test to be diagnostic of ability. Thus, I collapsed over the test description manipulation. Previous researchers have evoked stereotype threat in students by merely introducing the difficult test and without explicitly priming the negative stereotype (Spencer et al., 1999). Apparently, this stereotype threat evocation occurred in the current study.

Test Performance

Did the ethnicity of the test administrator differentially affect African American and European American students' performances? The pattern of means in Figure 4 suggests that it did.

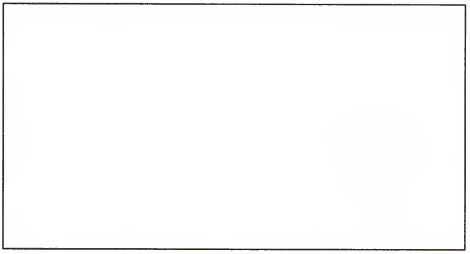

Figure 4. Mean test scores.

When the test administrator was African American, European American (M=32.1, SD=13.4) and African American students (M=32.1, SD=13.0) performed equally well on the test. However, when the administrator was European American, African American students performed worse on the test (M=22.1, SD=10.0) than did European American students (M=35.0, SD=13.2).

Statistical analysis using a two-way ANOVA (Student Ethnicity X Test Administrator Ethnicity) revealed a main effect for student ethnicity, $F(1, 111) = 7.00$, $p < .01$, $eta^2 = .06$, qualified by an interaction of student ethnicity and administrator ethnicity, $F(1, 111) = 6.97$, $p = .01$, $eta^2 = .06$. Planned contrasts showed that when the administrator was European American, African American students performed worse than European American students, $t(113) = 3.77$, $p < .001$, $eta^2 = .11$. When the administrator was African American, European American and African American students' performances did not differ, $t(113) = .23$, $p = .82$, $eta^2 = .00$. Further, African American students performed worse when the administrator was European American than when the

administrator was African American, t (113) = 2.52, $p < .05$. eta^2 = .05. European American students' performances were unaltered by the test administrator's ethnicity, t (113) = .901, p = .37, eta^2 = .00.

With the exception of students recruited over the phone, I identified interest in verbal-oriented course work for participants by their responses to an interest questionnaire administered during the experiment. The data for students who fell in the bottom tertile in terms of their interest in verbal-oriented course work were excluded from previous analyses. However, did these low interest students also show stereotype threat? Analyses of their performances revealed that they did not. Specifically, a 2 (Ethnicity of Student) X 2 (Ethnicity of Administrator) ANOVA of test performance revealed no main effects (both $Fs(1, 52) < 1.62$, $ps > .20$) or interaction ($F(1, 52) = 1.83$, $p = .18$) involving student and administrator ethnicity. Thus, consistent with previous evidence (Aronson et al., 1999), identification with the relevant domain appears to be a necessary condition for stereotype threat to hinder performance.

Did describing the test as diagnostic or nondiagnostic of intellectual performance affect students' performances? Analyses revealed no differences in test performance between the diagnostic (M=31.2, SD=13.0) and nondiagnostic conditions (M=30.8, SD=13.7), F (1, 113) = .024, p = .88, eta^2 = .00. In contrast to previous research (Steele & Aronson, 1995), test description did not affect performance. However, given that the manipulation checks revealed that all participants believed the test to be a diagnostic measure of intellectual ability, I concluded that students in both conditions experienced stereotype threat.

Anxiety

I predicted that African American students would report less anxiety when the test administrator was African American than when the administrator was European American. Consistent with prior studies (Aronson et al., 1999; Steele & Aronson, 1995), analyses revealed no main effects for student ethnicity, F (1, 111) = .933, p = .34, eta^2 = .00, or administrator ethnicity, F (1, 111) = 1.37, p = .24, eta^2 = .01, and no interaction between student ethnicity and administrator ethnicity, F (1, 111) = 1.16, p = .29, eta^2 = .01. However, the low Cronbach's alpha for Spielberger's State Anxiety Scale (α = .56) suggests that the scale may not be tapping a uni-dimensional construct.

Researchers have examined the impact of positively and negatively phrased items on the validity of anxiety scales ("I feel calm" versus "I feel anxious") and find that negating an item affects the students' responses to that item (Pilotte & Gable, 1990; Schotte, Maes, Cluydts, & Cosyns, 1996). Others suggest that Spielberger's State Anxiety Scale does not measure a uni-dimensional construct, because positive emotionality items such as "I am joyful" do not tap into the test taking experience (Sapp, Durand, & Farrell, 1995). Thus, items that test positive affect may be nonessential to investigations of stereotype threat.

Consequently, I removed all positive affect items from the Spielberger's State Anxiety Scale and summed the eight negative anxiety items (Cronbach's α = .83). The revised anxiety scale includes "I am tense," "I am regretful," "I feel upset," "I am presently worrying over possible misfortunes," "I feel anxious," "I feel nervous," "I am worried,"

and "I feel rattled." The pattern of anxiety scores are presented in Figure 5 and correspond to the pattern of test scores presented in Figure 5.

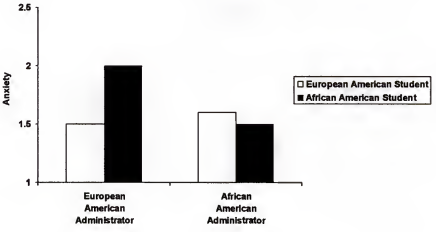

Figure 5. Mean responses to negative anxiety items.

When the test administrator was African American, European American (M=1.6, SD=.11) and African American students (M=1.5, SD=.11) reported equal levels of negative anxiety. However, when the administrator was European American, African American students reported more negative anxiety (M=2.0, SD=.12) than did European American students (M=1.5, SD=.09).

A Student Ethnicity by Test Administrator Ethnicity ANOVA on the negative anxiety scores revealed a significant interaction, $F(1, 111) = 5.15$, $p < .05$, $eta^2 = .04$. Contrasts showed that when the administrator was European American, African American students reported more negative anxiety than did European American students, $t(113) = 2.88$, $p < .01$, $eta^2 = .06$. When the administrator was African American, African American and European American students reported equal levels of negative anxiety, $t(113) = .321$, $p = .75$, $eta^2 = .00$. African American students reported more negative

anxiety when the administrator was European American than African American, t (113) = 2.74, $p < .01$, $eta^2 = .06$.

To examine whether participants' test scores were driven by negative anxiety, I included anxiety as a covariate (after centering; see Aiken & West, 1993) in the analysis of test scores using procedures recommended by Hull, Tedlie, and Lehn (1992). If the change in participants' test scores were driven by anxiety, then the significant Student Ethnicity by Administrator Ethnicity interaction, F (1, 111) = 6.97, $p = .01$, $eta^2 = .06$, would no longer be significant when the covariate was included in the analysis. Analyses revealed that the Student Ethnicity by Experimenter Ethnicity interaction for test scores was only marginally significant when anxiety was included as a covariate, F (1, 108) = 2.35, $p = .06$, $eta^2 = .02$. Thus, negative anxiety appears to partially mediate the interaction of Student Ethnicity and Test Administrator Ethnicity.

Performance Expectations

Researchers have theorized that performance expectations are related to test performance (Steele & Aronson, 1995). However, Aronson et al. (1999) and unpublished work reported in Steele and Aronson (1995) have not supported this assertion. I conducted an additional analysis to reexamine whether stereotype threatened and non-threatened students differ in performance expectations. As shown in Figure 6, when the test administrator was European American, European American students (M=60.4, SD=2.5) anticipated a higher performance than did African American students (M=51.6, SD=3.2). When the administrator was African American, African American students

(M=57.6, SD=2.9) and European American students (M=52.0, SD=2.9) anticipated performing relatively similar.

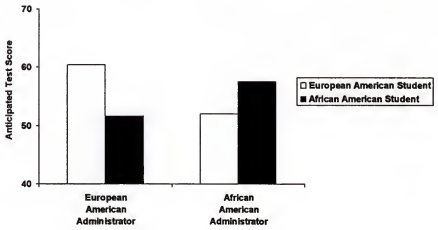

Figure 6. Mean anticipated test scores.

A two-way ANOVA revealed an interaction between Student Ethnicity and Test Administrator Ethnicity for expected test score, F (1, 111) = 6.25, p = .01, eta^2 = .05. Contrasts showed that when the administrator was European American, European American students anticipated a better performance than did African American students, t (113) = 2.15, $p < .05$, eta^2 = .04. However, when the administrator was African American, African American and European American students' anticipated scores did not differ significantly, t (113) = 1.38, p = .17, eta^2 = .02.

To examine whether participants' actual test scores were driven by their expected scores, I included expected score as a covariate in the analysis of test scores. If the students' test scores were due to the students' performance expectations, then the significant Student Ethnicity by Test Administrator Ethnicity interaction term, F (1, 111) = 6.25, p = .01, eta^2 = .05, would no longer be significant when expected score was included

in the model as a covariate. Analyses revealed that including expected score in the model produced no reduction in the significant Student Ethnicity by Test Administrator Ethnicity interaction for test scores, $F(1, 108) = 2.47$, $p < .05$, $eta^2 = .02$. Consistent with previous findings (Aronson et al, 1999; Steele & Aronson, 1995), performance expectations were unrelated to test performance.

Summary

These results suggest that the performance decrement among African American students found in prior research may be attributable to the ethnicity of the test administrator. When the test administrator was European American, then African American students performed poorly. When the test administrator was African American, then African American students and European American students performed equally well. European American students, by contrast, performed equally well regardless of whether the test administrator was European American or African American. Moreover, the anxiety reports corresponded to the test scores. African American students evaluated by a European American test administrator reported more anxiety than students in all other conditions. These findings provide preliminary support for the proposed model.

STUDY 2: OVERVIEW AND HYPOTHESES

Study 2 explored additional components of the proposed model of stereotype threat. Specifically, Study 2 examined whether self-presentational concerns led to distraction and then anxiety, and whether the ensuing attentional deficit facilitated performance on an easy task and hindered performance on a difficult task. As in Study 1, either an African American or European American test administrator informed students that they were taking a test that was diagnostic of intellectual ability. African American and European American students completed either a highly difficult test or a relatively easy test. During the test, students completed items that measured distraction, evaluation apprehension, self-consciousness, anxiety, and affect, followed by open-ended thought-listings. The Self-Consciousness Scale and the Positive and Negative Affect Scale were included in order to explore the participants' internal verses external orientation and affective state during the study. I imbedded an easy item at the beginning and end of the difficult test in an attempt to replicate Bond's (1982) finding that distraction increases during a test as students realize that they are missing items.

On completion of the test, participants completed an appraisal of the study and the test administrator. The appraisal included items assessing the extent to which the student perceived the administrator as someone who endorses cultural stereotypes.

When the test administrator was European American and the test was difficult, I predicted that African American students would perform worse on the test than would

European American students. I predicted that test administrator performance expectations, distraction, and anxiety would drive the decline in performance for African American students. Finally, I predicted that African American students would be more likely to miss the easy items imbedded at the end rather than at the beginning of the test.

When the test administrator was European American and the test was easy, I expected that African American students would perform better on the test compared with when the administrator was African American. Again, test administrator performance expectations, distraction, and anxiety would predict test performance for African American students.

When the test administrator was African American, I predicted that African American students and European American students would perform equally well on the easy and difficult tests. I also predicted in this condition that test administrator performance expectations, distraction, and anxiety would not predict test performance.

Finally, I predicted that the performances of European American students would not be affected by the administrator's ethnicity, administrator performance expectations, distraction or anxiety.

Method

Participants

Fifty-eight European American and 48 African American students participated in exchange for partial course credit. All undergraduate Introductory psychology students reported their approximate verbal scores on the Scholastic Aptitude Test (SAT) and their reported interest in verbal-oriented course work during mass prescreening (see Appendix

D). Students who scored in the national top tertile of verbal SAT scores (> 550) and in the top tertile of the interest item were contacted by phone and invited to participate.

The design was a 2 (Student ethnicity: European American vs. African American) X 2 (Test administrator ethnicity: European American vs. African American) X 2 (Task: Easy vs. Difficult) between-subjects factorial.

Procedure

Up to six participants arrived at the lab per session. Each student sat in separate cubicles or at a large table. Either an African American or a European American female test administrator informed the participants that they would be taking a 30-item test. The actual test contained only 20 items, however upon completion of the test, students were to believe that 10 items remained.

Prior to beginning the test, participants recorded their anticipated test score and also the score that they believed that the test administrator expected them to make (Appendix E). Participants placed their estimates in an envelope so that they would remain confidential and out of the administrator's view.

Participants completed a 20-item test that was either extremely difficult (see Appendix F) or not (see Appendix G). For the difficult test, items were extracted from the most difficult items of the verbal section of the Graduate Record Examination, the same items used by Steele and Aronson (1995) and in Study 1 of this dissertation. For the easy test, items were questions drawn from the Scholastic Aptitude Test preparation guide. On the difficult test, items three and seventeen were the same as those from the easy test. Bond (1982) found that students who were distracted were more likely to miss easy items

imbedded in a difficult test. Presumably, participants taking the difficult test would grow increasingly anxious and distracted as the test progressed, and consequently, would be more likely to miss the item embedded at the end of the test than the item embedded at the beginning of the test.

After students completed the test, the test administrator led the students to believe that they had only completed the first section of the test. They then responded to the Cognitive Interference Questionnaire (CIQ; Sarason, 1980; see Appendix H), the brief version of the Fear of Negative Evaluation Scale (Leary, 1983; see Appendix I), the Self-Consciousness Scale (Fenigstein, Scheier, & Buss, 1975; see Appendix J), eight negative items extracted from the state form of the State-Trait Anxiety Inventory (Spielberger et al., 1970; see Appendix K), the Positive and Negative Affect Scale (Watson, Clark, & Tellegen, 1988; see Appendix L), and an open-ended thought-listing question (Appendix M).

Past studies have collected psychological measures either prior to the test or after completion of the test (Steele & Aronson, 1995; Spencer et al., 1999; Aronson et al., 1999). Distraction and apprehension may have either subsided or not yet appeared at these points in the session. By collecting these measures during the test, I hoped to create a more sensitive measure of what participants were thinking while taking the test.

On completion of the questionnaires and open-ended thought-listings, the test administrator informed participants that the study was over. The administrator explained that the Institutional Review Committee at the University of Florida was interested in their experience as participants in this study. The administrator left the room, and the students

responded to a variety of questions about the administrator (Appendix N). The questionnaire included items inquiring about the test administrator's tendency toward racial and gender bias, and the student's interest in having the administrator as a mentor in the future. Participants placed the completed evaluation form in the envelope with the previous request for anticipated scores. They sealed the envelope and placed it in a box marked "Institutional Review Committee." The test administrator returned to the room and fully debriefed participants.

Results

Test Performance

Did the test performances of European American and African American students depend on test difficulty and test administrator ethnicity? Analysis of test performance using a three-way ANOVA (Student Ethnicity X Test Administrator Ethnicity X Test Difficulty) did not reveal the predicted three-way interaction, $F(1, 98) = .71$, $p = .40$, $eta^2 = .00$.

I predicted that the presence of a European American test administrator would facilitate performance for the African American students on the easy test. However, examination of the test scores reveals that the easy test was not as easy as intended, certainly not easy enough to elicit a social facilitation effect. Instead, examination of the pattern of means in Table 1 reveals that the easy test was actually moderately difficult ($M = 78.1$, $SD = 14$), while the difficult test was extremely difficult ($M = 38.3$, $SD = 12.6$).

Table 1.
Grade on test (% correct) as a function of Test Administrator Ethnicity, Student Ethnicity, & Test Difficulty.

	Test Difficulty			
	Easy		Difficult	
	EA Administrator	AA Administrator	EA Administrator	AA Administrator
EA Student	87.0 (9.78)	88.2 (15.5)	43.8 (11.8)	37.9 (9.41)
AA Student	72.5 (19.6)	78.3 (11.6)	30.8 (15.8)	38.3 (9.85)

Note. EA = European American, AA = African American

Analysis did, however, reveal a significant main effect of student ethnicity, $F(1,98)=12.70$, $p<.01$, $eta^2=.11$, and a marginally significant interaction of student ethnicity and test administrator ethnicity, $F(1, 98) = 3.02$, $p = .09$, $eta^2=.03$. While the results from the two-way interaction using the omnibus ANOVA fell short of conventional levels of significance, the pattern of means (see Table 1) was just as predicted and consistent with Study 1. Moreover, the results of the a priori comparisons using the pooled error term (MSE = 1.75) revealed the predicted effects (see Figure 7). Overall, African America students ($M=51.7$, $SD=27.5$) performed worse than European American students ($M=64.1$, $SD=24.4$) when the test administrator was European American, $t(102)=1.77$, $p=.07$, $eta^2=.03$. When the test administrator was African American, the performance of African American students ($M=58.3$, $SD=23.0$) and European American students ($M=65.0$, $SD=28.6$) did not differ, $t(102)=.910$, $p=.37$, $eta^2=.00$.

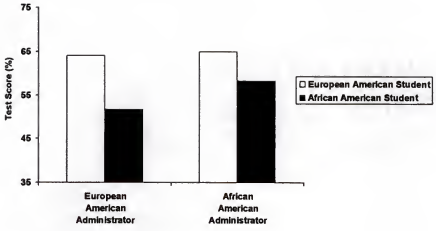

Figure 7. Test scores collapsed over test difficulty.

As shown in Figure 8, this pattern of means was particularly strong for the students who took the difficult test, which was almost identical to the test used in Study 1. When the test administrator was African American, European American and African American students performed equally well on the test. However, when the administrator was European American, African American students performed worse on the test than did European American students.

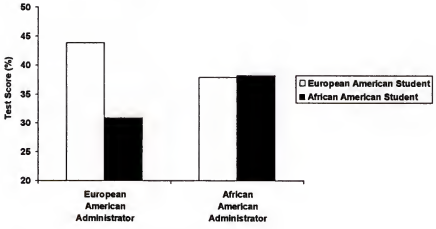

Figure 8. Test scores on the difficult test.

On the difficult test, statistical analysis using a two-way ANOVA (Student Ethnicity X Test Administrator Ethnicity) revealed a main effect for student ethnicity, $F(1, 51) = 4.08$, $p < .05$, $eta^2 = .07$, qualified by an interaction of student ethnicity and administrator ethnicity, $F(1, 49) = 4.07$, $p < .05$, $eta^2 = .08$. Planned contrasts showed that when the test administrator was European American, African American students (M=30.8, SD=15.8) performed worse than European American students (M=43.8, SD=11.8), $t(105) = 2.56$, $p = .01$, $eta^2 = .06$. When the test administrator was African American, European American (M=37.9, SD=9.4) and African American students (M=38.3, SD=9.9) performed similarly, $t(105) = .00$, $p = .97$, $eta^2 = .00$.

On the easy test, analyses of test performance revealed a main effect for student ethnicity, $F(1, 51) = 8.73$, $p < .01$, $eta^2 = .15$, and no interaction of student ethnicity and administrator ethnicity, $F(1, 49) = .429$, $p = .52$, $eta^2 = .00$. However, planned contrasts showed that when the test administrator was European American, African American students (M=72.5, SD=19.6) performed worse than European American students, (M=87.0, SD=9.8), $t(105) = 2.83$, $p < .01$, $eta^2 = .07$ (see Figure 9). When the test administrator was African American, African American (M = 78.3, SD = 11.6) and European American (M = 88.2, SD = 15.5) students performed similarly, $t(105) = .910$, $p = .37$, $eta^2 = .00$. In sum, the pattern of means for the difficult and easy test parallel the pattern of means found in Study 1. These findings support my primary hypothesis that African American students would perform worse than European American students on a difficult test only when the test was administered by a European American.

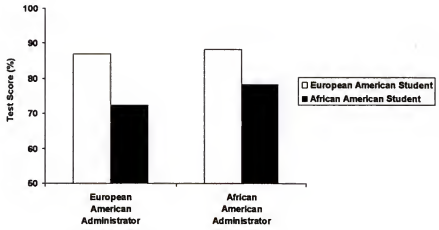

Figure 9. Test scores on the easy test.

Exploring the Source of the Poor Test Performance Among African American Students

I predicted that test administrator ethnicity would predict perceptions of the test administrator's performance expectations. Specifically, African American students would anticipate lower performance expectations from the European American test administrator due to the administrator's implicit endorsement of stereotypes about African American intellectual capabilities. Administrator performance expectations would create interfering thoughts, which in turn would produce anxiety. The anxiety would then culminate in poor performance. I collected measures of perceptions of test administrator performance expectations prior to taking the test, and anxiety and distraction at a point participants believed was midway through the test (see Table 2 for means and standard deviations). Test administrator performance expectations constituted a single-item measure of the students' estimate of the score that they believed that the test administrator expected them to make.[1]

[1] I also assessed evaluation apprehension and the extent to which participants perceived the test administrator as biased and prejudiced. However, the measures of evaluation apprehension (α = .69) and perceived test administrator bias were unreliable (α = .23), and thus, I was forced to exclude them from the analysis.

Table 2.
Proposed mediators as a function of test administrator ethnicity and student ethnicity. Means are collapsed over test difficulty.

	Test Administrator Ethnicity							
	European American				African American			
	EA Student		AA Student		EA Student		AA Student	
	M	SD	M	SD	M	SD	M	SD
Test Administrator Performance Expectations	46.7	(10.4)	43.5	(12.7)	52.2	(7.0)	54.0	(17.0)
Distraction	2.24	(.54)	2.51	(.66)	2.23	(.42)	2.46	(.56)
Anxiety	1.35	(.42)	1.50	(.42)	1.52	(.58)	1.43	(.52)

Although research has found no evidence for distraction, researchers maintain that stereotype threat likely hinders performance through inefficient information processing (Steele & Aronson, 1995). Thus, I added a second measure of distraction to Study 2. Previous work suggests that students will become progressively distracted when completing a difficult test in the presence of an evaluative audience (Bond, 1982). Such students are likely to miss easy items embedded at the end of a difficult test due to the heightened distraction. As an additional measure of distraction, I embedded two easy items in the tests: one at the beginning of the test (item # 3) and one at the end of the test (item # 17). I predicted no differences in performance on the first easy item as a function of student and test administrator ethnicity. However, I predicted that African American students would be more likely than European American students to miss the second easy item embedded in the tests, but only when the test administrator was European American. The pattern of means in Figures 11 and 12 support this prediction. All students were

equally likely to perform well on the item embedded at the beginning of the test. However, African American students were more likely than European American students to miss the item embedded at the end of the test when the test administrator was European American. When the test administrator was African American, an equal percentage of European American and African American students responded correctly to the second embedded item.

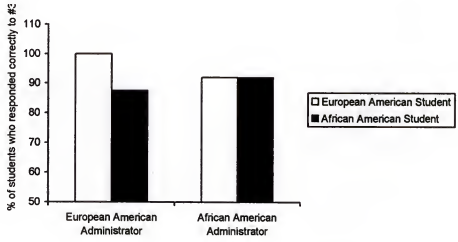

Figure 11. Percentage of students who responded correctly to the first embedded item.

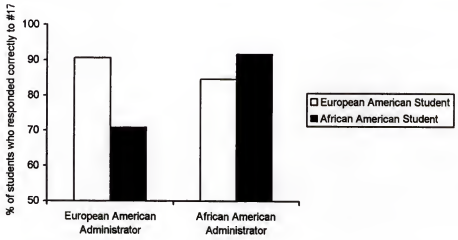

Figure 12. Percentage of students who responded correctly to the second embedded item.

On the first easy, embedded item, statistical analyses used Yates' Correction for Continuity as a means to correct for the small frequency of incorrect responses on the item (Howell, 1982). Analysis of the first item revealed that European American and African American students did equally well regardless of whether the test administrator was European American, $\chi 2$ $(1, \underline{N} = 106) = 2.07$, $\underline{p} > .05$, or African American, $\chi 2$ $(1, \underline{N} = 106) < 1$, $\underline{p} > .05$. However, analysis of the second easy, embedded item revealed that African American students were more likely than European American students to miss the item when the test administrator was European American, $\chi 2$ $(1, \underline{N} = 106) = 3.66$, $\underline{p} < .06$. When the test administrator was European American, 91% of the European American students and only 71% of the African American students responded correctly to the item. However, when the test administrator was African American, European American and African American students were equally likely to respond correctly to the second easy, embedded item, $\chi 2$ $(1, \underline{N} = 106) < 1$, $\underline{p} > .05$.

Path Analyses

I used a recursive (unidirectional) path analysis model to investigate the hypothesis that test administrator ethnicity would influence test performance for African American students through the effects of administrator expectations, anxiety, and distraction (see Figure 10).

Because European American students test performances do not vary with the test administrator, I did the path analysis only on African American students. The correlation matrix for each of the variables used in the path analysis is presented in Table 3. The path analysis consisted of four regressions, with direct effects measured on test performance in

the final regression and indirect effects assessed in the first three regressions. Beta weights representing the standardized coefficients are presented in Figure 10.

Table 3.
Correlations among proposed mediators for the Test Administrator Ethnicity x Student Ethnicity interaction for African American students.

	Administrator Ethnicity	Administrator Performance Expectations	Distraction	Anxiety	Test Score
Administrator Ethnicity	--				
Administrator Performance Expectations	.34*	--			
Distraction	-.05	-.29*	$\alpha = .82$		
Anxiety	-.07	.03	.30*	$\alpha = .81$	
Test Score	.13	-.01	-.34*	-.38**	$\alpha = .88$

* $p \leq .05$, ** $p \leq .01$.
Note. Test Administrator ethnicity and administrator performance expectations were single item measures. The European American test administrator was coded as 0 and the African American test administrator was coded as 1.

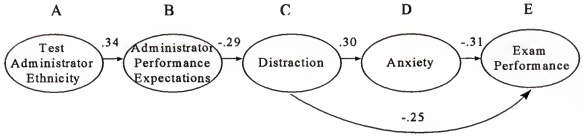

Figure 10. Path model and standardized coefficients.

In the first regression, I tested whether African American students assumed that a European American administrator would expect them to do poorly on the test. I predicted that test administrator ethnicity would predict perceived test administrator performance expectations (see Link A → B). Regression analyses revealed that test administrator ethnicity did predict perceived test administrator performance expectations, $\beta = .34$, $p < .05$. Next, I anticipated that perceived administrator performance expectations would predict distraction (see Link B → C). In the second regression sequence, analyses revealed that perceived administrator performance expectations predicted distraction, $\beta = -.29$, $p = .05$.

In the third regression analysis, I tested whether distraction predicted anxiety (see Link C → D). Analyses revealed that distraction was a significant predictor of anxiety, $\beta = .30$, $p < .05$. In the final direct analysis, I tested whether anxiety predicted test performance for African American students. I found support for a direct effect of anxiety on test performance (see Link D → E), $\beta = -.34$, $p < .05$. Distraction was also a significant predictor of test performance (see Link C → E), $\beta = -.34$, $p < .05$.

How well did the model presented in Figure 10 fit the data? Multiple indices can describe the fit of the model to the data. The chi-square statistic was not significant, χ^2 (5, $n = 48$) = 2.86, $p = .72$. The small chi-square suggests that the model is sufficiently complex. Multiple criteria measured the goodness-of-fit of the model considering a 1.0 to be perfect and a 0 to indicate that the model is no better fit than the null hypothesis. To accommodate the small sample size, I reported the Comparative Fit Index which indicated a good fit, CFI = 1.00. Further, the Root Mean Square Residual (RMR = .05) and the

Root Mean Square Error of Approximation (RMSEA = .00) confirmed that the fit is adequate (an inadequate fit >.05). According to these multiple criterion, I can conclude that the data adequately fits the proposed model. The model is both recursive and consistent with the theoretical hypotheses.

Next, I tested the assumption that the model would not fit equally well for European American students. I included both groups in the analysis, and first, tested the fit of a model in which both groups' data loaded onto each factor in the model equally well. To do so, I constrained the factor loadings to be equal across groups, χ^2 (15, n = 106) = 10.10, p = .81. I compared this chi square to a second analysis in which I tested the fit of a model in which both groups' data varied across factor loadings, χ^2 (10, n = 106) = 6.49, p = .77. Allowing the factor loadings to vary across groups led to a smaller chi square and better fit indicating that there is not equal accuracy when testing European American and African American students on the same dimensions.

Further, examination of specific factor loadings and standard errors for the two groups (see Table 4) indicate that, consistent with the proposed model, paths from administrator performance expectations to distraction, anxiety to grade, and distraction to grade describe the data for African American students more so than for European American students. Most notably, for European American students, test administrator performance expectations are not related to distraction, and distraction was unrelated to performance.

Table 4.
Factor loadings and standard errors.

	African American Students					European American Students				
	A	B	C	D	E	A	B	C	D	E
A	--	--	--	--	--	--	--	--	--	--
B	.34	--	--	--	--	.29	--	--	--	--
	(.14)					(.13)				
C	--	-.29	--	--	--	--	-.01	--	--	--
		(.14)					(.13)			
D	--	--	.30	--	--	--	--	.27	--	--
			(.14)					(.13)		
E	--	--	-.25	-.31	--	--	--	-.09	-.20	--
			(.14)	(.14)				(.13)	(.13)	

Note. A = Test administrator ethnicity, B = Administrator performance expectations, C = Distraction, D = Anxiety, E = Test performance

Summary

These results corroborate the findings of Study 1 and provide additional insight into the processes leading to the performance decrement among stereotype threatened African American students. As in Study 1, when the test administrator was European American, African American students performed worse than European American students. When the test administrator was African American, African American students and European American students performed equally well. The effect of test administrator ethnicity on student performance was driven by the students' perceptions that the outgroup test administrator held low expectations for their test performance, and by distraction and anxiety. The performances of European American students, by contrast, seem relatively unaffected by the ethnicity of the test administrator. Further, administrator expectations predicted distraction, distraction led to anxiety, and both anxiety and distraction led directly to test performance. This model appears to represent an adequate

account of the process by which the ethnicity of the test administrator ultimately affects performance on tests.

Ancillary Hypotheses

Analyses have supported the hypothesis that stereotype threat involves interfering thoughts during the test. However, I collected additional student self-report measures in an effort to investigate the attentional focus of these thoughts. Previous researchers have speculated that stereotype threat involves a fear of confirming a stereotype both to one's self and to others (Steele & Aronson, 1995).

I have argued that stereotype threat involves a fear of being unable to refute an inaccurate stereotype about one's group to an evaluative audience. Accordingly, stereotype threatened students perform poorly on a test when the test administrator is European American because they are distracted by their efforts to dispel a negative stereotype to an external audience. Prior researchers, however, have proposed that students fear confirming a stereotype to one's self, which assumes that students are engaging in self-focused rumination during the test.

To test this alternative hypothesis about the thoughts of threatened students, I analyzed students' responses on measures of public and private self-consciousness (Fenigstein, Scheier, & Buss, 1975). The private self-consciousness scale assesses a person's tendency toward private self-focus. The public self-consciousness scale assesses a person's concern over the self within social contexts. While the private and public self-consciousness scales do not measure self-focused rumination and self-presentational concerns per se, they seem adequate proxies for the constructs.

A 2 (Student Ethnicity) X 2 (Test Administrator Ethnicity) ANOVA on participants' responses to the public self-consciousness scale revealed no significant main effects or interaction (all Fs (1,102) < 1.20, ps > .20). Thus, there is no difference across student ethnicity or test administrator ethnicity on public self-consciousness. However, consistent with my predictions, the correlations in Table 5 show a negative relationship between public self-consciousness and test performance among African American students ($r = -.28$, $p = .05$), but not among European American students ($r = -.02$, $p = .91$). The more African American students focused on their public image, the more their test performance faltered. However, public self-consciousness was unrelated to test performances among European American students.

Table 5.
Correlations between measures and grade on test as a function of student ethnicity.

	Student Ethnicity	
Measure	EA Students (n = 58)	AA Students (n = 48)
Administrator Performance Expectations	-.08	-.01
Distraction	-.15	-.35**
Anxiety	-.23	-.39**
Public Self-Consciousness	-.02	-.28*
Private Self-Consciousness	-.03	-.03

* $p \leq .05$, ** $p \leq .01$.

Analysis of participants' responses to the measure of private self-consciousness revealed only one effect: a main effect for ethnicity, $F(1, 104) = 3.97$, $p < .05$, $eta^2 = .04$. However, the pattern of means were contrary to the hypotheses of previous researchers that African American students are engaging in greater rumination than European

American students. Specifically, European American students ($\underline{M}$ = 3.52, $\underline{SD}$ = .60) reported more private self-consciousness during the test than did African American students ($\underline{M}$ = 3.28, $\underline{SD}$ = .61).

These findings, while based on proxy measures of self-rumination and public self-focus, are nevertheless consistent with the assertion that a focus of attention on external evaluation and not a self-focused rumination is responsible for the poor performances of African American students. Whereas greater pubic self-consciousness was associated with poorer performances among African American students, it was not associated with poorer performances among European American students. Moreover, African American students were less privately self-focused, suggesting less self-rumination and presumably self-doubt than found among European American students.

Discussion

The present studies suggest a process by which stereotype threat impairs performance. Prior researchers (Steele & Aronson, 1995) have hypothesized that stereotype threat impairs performance by way of self-doubt and fear of confirming a stereotype to one's self. I proposed, however, that the performance decline is caused by a desire to convey competency to an evaluative audience who is presumed unlikely to accept this image. This predicament produces distraction and anxiety, which in turn impair performance. I further assumed that African American students would be more likely than European American students to believe that a European American test administrator would hold a negative stereotype of them. The results would be lower test performance among African American students confronting a European American test administrator

than African American students confronting an African American test administrator. Consistent with my predictions, in both Studies 1 and 2, African American students performed worse than European American students on a difficult test only when the test administrator was European American. European American and African American students performed equally well when the test administrator was African American.

Study 1 provided evidence that the effect of test administrator ethnicity on test performance among African Americans was partially mediated by anxiety. Study 2 replicated the anxiety effects of Study 1 and further found that the effect of test administrator ethnicity on performance was also mediated by perceived test administrator performance expectations, and by distraction. Specifically, the results revealed that African American students paired with a European American test administrator perceived that the administrator would evaluate them according to negative group stereotypes (i.e., low intelligence). These perceptions were linked to greater distracting, interfering thoughts, diverting valuable cognitive resources from performing the test. The interfering thoughts were associated with poorer test performance both directly and indirectly through increased anxiety.

Study 2 provided additional evidence that African American students paired with a European American administrator were distracted while performing the task. Specifically, African American and European American students performed equally well on an easy item embedded at the beginning of the difficult test. However, when paired with an African American test administrator, African American students were more likely than European American students to miss the easy, embedded item at the end of the test.

Implications for Research on Stereotype Threat

The present findings extend our understanding of stereotype threat in several ways. First, test administrator ethnicity appears to be an important condition for stereotype threat to hinder performance. This criterion has eluded past research findings, because most test administrators have not shared the students' group membership. Second, the present research provides evidence for the process by which stereotype threat hinders performance. My findings suggest that African American students perform worse on an intelligence test than similarly capable European American students because African American students are more distracted and anxious than European American students and the greater distraction and anxiety comes from the perception that their test administrator evaluates them negatively.

Third, the present findings suggest that the poorer performance of African American students arises from concern about confirming a negative stereotype to an external audience, but not to themselves. Stated otherwise, African American students perform poorly because they are focused on dismissing a negative audience expectation and not because they doubt their ability (self-doubt). Three pieces of evidence suggest that self-doubt is not producing the effect. First, if self-doubt was driving the performance decline, as speculated by Steele and Aronson (1995), ethnicity of the test administrator should not be a necessary condition for the performance decline. Presumably, after the stereotype is made salient, self-doubt would endure despite variations in the external setting such as the test administrator's ethnicity. Second, prior research reveals that stereotype threat hinders performance only among students who report high interest and

ability in the performance domain (Aronson, et al., 1999). If the performance decline depends on self-doubt, presumably less assured students would be more susceptible to the effects of stereotype threat. Third, the proxy measures of self-rumination and public self-focus reported in Study 2 revealed a pattern inconsistent with self-doubt. Specifically, the more African American students focused on their public image, the more their test performance faltered. Further, European American students reported more self-focused attention than did African American students. Taken together, these findings suggest that stereotype threat does not arise from self-doubt. The evidence suggests instead that stereotype threat arises from one's concerns about an evaluator's implicit rejection of one's capacity to demonstrate an important ability.

This work also sheds light on the role of distraction in stereotype threat. Past studies have failed to find evidence for distraction, but researchers have maintained that distraction likely plays some role in the performance decline (Steele & Aronson, 1995). This research supported a clear link between perceptions of the European American test administrator and distraction, and between distraction, anxiety and test performance. In past research, students completed self-report measures of distraction after completing the test. The timing, however, might have been unwise in that after the test, feelings of distraction may have subsided. I examined distraction in two ways: first, by embedding easy items in the difficult test, and second, by collecting a self-report measure of distraction at a point where the students believed the test was only half over. Thus, the present research examined interfering thought processes as they occurred during the test.

Finally, while previous research has yielded mixed support for anxiety in stereotype threat, this research found evidence in two studies that anxiety plays a role in stereotype threat. In Study 1, the pattern of means for negative anxiety mapped onto the pattern of performance outcomes and partially mediated the effects of test administrator ethnicity on performance. African American students reported more anxiety than did European American students when the test administrator was European American, and the greater anxiety appears partially responsible for their subsequent performance decline. Study 2 provided evidence that distraction is associated with heightened anxiety, and anxiety is associated with a performance decline. Distraction also influences performance independently of anxiety. This finding is consistent with the hypothesis that anxiety increases as students became aware of their poor performance caused by the onset of interfering thoughts. In past studies, students completed items measuring state anxiety at the end of the test when, like distraction, anxiety may have diminished. I collected a more specific measure of state anxiety (including only negative statements about anxiety) at a point where students believed that the test was only half over.

Importantly, my model provides only preliminary information about the processes that underlie stereotype threat and its consequences. Future investigations should isolate the key factors in the model to expand our understanding of their impact on performance. For instance, in Study 2, test administrator performance expectations induced distraction. A replication might manipulate test administrator performance expectations directly in order to understand the limits and conditions of its impact on distraction. In addition, researchers should operationalize distraction according to methods used in previous work.

For instance, test performance should decline for African American students with the introduction of intrusive stimuli (i.e., sounds and repetitive tasks) if they are, in fact, experiencing overtaxed attention. Isolating cognitive load would verify its role in the performance decline and provide information about the amount and type of cognitions that constitute overload during performance tasks.

Future investigations might also continue to explore the possibility of a social facilitation effect with stereotype threat. In Study 2, I manipulated test difficulty to explore possible facilitation effects caused by stereotype threat. Drawing from research on social facilitation, I speculated that arousal due to the European American test administrator would facilitate the performances of African American students on the easy test (i.e., a well-learned task). However, test difficulty had no effect on performance outcomes. When the test administrator was European American, African American students performed worse than European American students on the easy and the difficult tests. Apparently, even the easy test was not easy, but rather moderately difficult leaving little opportunity for a facilitation effect. To examine performance facilitation in stereotype threat, it will be necessary to develop a much easier test. To compensate for the fact that all students will achieve perfect scores, speed rather than accuracy may be a better measure. In addition, researchers may look at tasks that require automatic processing as opposed to varying degrees of high cognitive effort.

It would be of interest to the current theoretical model and real-world applications of stereotype threat to find evidence for a facilitation effect. Given that attentional explanations of social facilitation maintain that cognitive overload mediates the effect of an

evaluative audience on performance (Baron, 1986), a facilitated performance on an easier task would further corroborate the role of distraction and audience apprehension in stereotype threat. Further, any improvement in performance due to stereotype threat would support my contention that threatened students experience heightened motivation as opposed to self-doubt at the onset of the test. When the demands of the task allow, motivation should reveal itself in an improved performance.

It will also be important to test whether this model applies to investigations of stereotype threat with non-cognitive performance outcomes. Studies have demonstrated the performance decrements on non-cognitive tasks such as sports ability (Stone, Lynch, Sjomeling, & Darley, 1999). Given that the crux of the current model is cognitive overload, different processes may account for the effect of stereotype threat on non-cognitive tasks.

Past studies have described the impact of "solo status" within the classroom. Specifically, one's ethnicity or gender may interfere with information processing when a trait differs from others in the setting (Inzlicht & Ben-Zeev, 2000). Thus, an African American woman might experience interference in an all-male classroom even if the males were a mix of European American and African American students, but not in a classroom comprised of African American male and female students. It would follow that future investigations might explore the composition of the classroom in addition to the test administrator to explore ways in which these factors may interact to affect performance.

It is also crucial that we understand whether the current model can be generalized to other student populations. Based on the present model and findings, anyone who

belongs to a group to which a negative stereotype about intellectual ability applies will be susceptible to the threat and subsequent performance impairment. Such groups might include the elderly, students from lower socioeconomic backgrounds, Hispanic students, and women. People need merely be aware of the stereotype to perceive that the test administrator adheres to it. The model may explain the finding that even European American men are susceptible to the adverse effects of stereotype threat (Aronson, et al., 1999). Possibly, a cultural stereotype does not need to exist. Rather, all that is necessary is that the test administrators explicitly express low expectations for the group to which the students belong. Presumably, any students may be susceptible to stereotype threat if the administrators convey low confidence in the students' abilities either implicitly by being a member of a privileged group or explicitly through their comments and actions.

For particular groups, stereotype threat may impact performance in a different way. For instance, students who belong to groups with positive academic stereotypes (e.g., Asian Americans) may face a new and separate set of responses to perceived test administrator expectations. Recent research reveals that Asian American women perform poorly on a math test when the public's high performance expectations for them are made salient (Cheryan & Bodenhausen, 2000). Further, women in the classroom, particularly in math and science courses, may present an exception to this model. As mentioned previously, the pervasiveness of female teachers and test administrators may lead female students to assume that their expectations will reflect those of the rest of society.

Implications for the Educational Setting

This research poses critical questions about the testing environment in schools and testing centers. While additional research is needed, this work warrants consideration of matching instructor and student ethnicity within educational settings. Such a need would warrant support of initiatives geared toward ensuring adequate representation of the nation's diversity within educational institutions. Considering the present research, such programs may enhance standardized test scores and class performances of traditionally underrepresented groups.

This research may merit sensitizing teachers and test administrators to the impact of their implicit messages about students' abilities and potential. Further, teachers and test administrators may be trained to convey appropriate expectations explicitly to combat implicit messages that the may inadvertently send to students. Future studies might tease out the most appropriate message and effective mode of transmitting the message. Ultimately, teachers and administrators must ensure that all students feel that their ability and potential is equally recognized within classrooms and testing environments.

Experts in education reform have speculated numerous causes of the gap between African American and European American students in the United States in standardized test scores. The current research offers a relatively simple explanation for the effect and suggests a possible way to narrow the performance gap between different groups of students. Group membership and perceived expectations of the test administrator must be addressed and accounted for in the testing environment. If standardized tests continue to be an important criterion in the selection of students to continue into higher education, it is

imperative that we isolate factors that lead to misrepresentation of the abilities and potential of students.

APPENDIX A
STUDY 1 PARTICIPANT SELECTION

Circle your current year in school:

Freshman 2. Sophomore 3. Junior 4. Senior

Major _____

Please circle below your typical attitude toward verbal-oriented classes (i.e., literature, languages, writing).

1	2	3	4	5	6	7
Strongly Dislike						Strongly Like

APPENDIX B
STUDY 1 TEST

Directions: Each sentence below has one or two blanks, each blank indicating that something has been omitted. Beneath the sentence are five lettered words or sets of words. Choose the word or set of words for each blank that best fits the meaning of the sentence as a whole.

1. The new biological psychiatry does not deny the contributing roles of psychological factors in mental illness, but that these factors may act as a catalyst on existing physiological conditions and _____ such illnesses.
A. disguise
B. impede
C. constrain
D. precipitate
E. consummate

2. During periods of social and cultural stability, many art academies are so firmly controlled by _____ that all real creative work must be done by the _____.
A. dogmatists . . disenfranchised
B. managers . . reactionaries
C. reformers . . dissatisfied
D. impostors . . academicians
E. specialists . .elite

3. The first World War began in a context of jargon and verbal delicacy and continued in a cloud of _____ as _____ as language and literature, skillfully used, could make it.
A. circumlocution . . literal
B. cliché . . lucid
C. euphemism . . impenetrable
D. particularity . . deliberate
E. subjectivity . . enthralling

4. While Parker is very outspoken on issues she cares about, she is not _____; she concedes the _____ of opposing arguments when they expose weaknesses in her own.
A. fickle . . validity
B. arrogant . . restraint
C. fanatical . . strength
D. congenial . . incompatibility
E. unyielding . . speciousness

5. Usually the first to spot data that were inconsistent with other findings, in this particular experiment she let a number of _____ results slip by.
A. inaccurate
B. verifiable
C. redundant
D. salient
E. anomalous

6. Cèzanne's delicate watercolor sketches often served as _____ of a subject, a way of gathering fuller knowledge before the artist's final engagement of the subject in an oil painting.
A. an abstraction
B. an enchantment
C. a synthesis
D. a reconnaissance
E. a transcription

7. Liberty is not easy, but far better to be an _____ fox, hungry and threatened on its hill, than a _____ canary, safe and secure in its cage.
A. unfriendly . . fragile
B. aging . . young
C. angry . . content
D. imperious . . lethargic
E. unfettered . . well-fed

8. Johnson never _____ to ignore the standards of decent conduct mandated by company policy if _____ compliance with instructions from his superiors enabled him to do so, whatever the effects on his subordinates.
A. deigned . . tacit
B. attempted . . helf-hearted
C. intended . . direct
D. scrupled . . literal
E. wished . . feigned

Directions: In each of the following questions, a related pair of words or phrases is followed by five lettered pairs of words or phrases. Select the lettered pair that best expresses a relationship similar to that expressed in the original pair.

9. didactic : instruct ::
A. pedantic : contend
B. comic : amuse
C. theatrical : applaud
D. imperative : obey
E. rhetorical : recite

10. garrulous : : talkative ::
A. suspicious : unreliable
B. cantankerous : obtuse
C. cloying : sweet
D. reflective : insightful
E. prudent : indecisive

11. digressive : statement ::
A. connotative : definition
B. slanderous : slur
C. tangential : presupposition
D. biased : opinion
E. circuitous : route

12. chicanery : clever ::
A. expertise : knowledgeable
B. certainty : doubtful
C. gullibility : skeptical
D. machination : heedless
E. tactlessness : truthful

13. pluck : quit ::
A. verve : flinch
B. gall : skimp
C. pride : grovel
D. charm : smile
E. poise : waver

The passage below is followed by questions based on its content. After reading the passage, choose the best answer to each question. Answer all questions following the passage on the basis of what is stated or implied in that passage.

The historian Frederick J. Turner wrote in the 1890's that the agrarian discontent that had been developing steadily in the United States since about 1870 had been precipitated by the closing of the internal frontier—that is, the depletion of available new land needed for further expansion of the American farming system. Not only was Turner's thesis influential at the time, it was later adopted and elaborated by other scholars, such as John D. Hicks in *The Populist Revolt* (1931). Actually, however, new lands were taken up for farming in the United States throughout and beyond the nineteenth century. In the 1890s, when agrarian discontent had become most acute, 1,100,000 new farms were settled, which was 500,000 more than had been settled during the previous decade. After 1890, under the terms of the Homestead Act and its successors, more new land was taken up for farming than had been taken up for this purpose in the United States up until that time. It is true that a high proportion of the newly farmed land was suitable only for grazing and dry farming, but agricultural practices had become sufficiently advanced to

make it possible to increase the profitability of farming by utilizing even these relatively barren lands.

The emphasis given by both scholars and statesmen to the presumed disappearance of the American frontier helped to obscure the great importance of changes in the conditions and consequences of international trade that occurred during the second half of the nineteenth century. In 1869 the Suez Canal was opened and the first transcontinental railroad in the United States was completed. An extensive network of telegraph and telephone communications was spun: Europe was connected by submarine cable with the United States in 1866 and with South America in 1874. By about 1870 improvements in agricultural technology made possible the full exploitation of areas that were most suitable for extensive farming on a mechanized basis. Huge tracts of land were being settled and famed in Argentina, Australia, Canada, and in the American West, and these areas were joined with one another and with the countries of Europe into an interdependent market system. As a consequence, agrarian depressions no longer were local or national in scope, and they struck several nations whose internal frontiers had not vanished or were not about to vanish. Between the early 1870s and the 1890s, the mounting agrarian discontent in America paralleled the almost uninterrupted decline in the prices of American agricultural products on foreign markets. Those staple-growing farmers in the United States who exhibited the greatest discontent were those who had become most dependent on foreign markets for the sale of their products. Insofar as Americans had been deterred from taking up new land for farming, it was because market conditions had made this period a perilous time in which to do so.

14. According to the author, changes in the conditions of international trade resulted in an

A. underestimation of the amount of new land that was being farmed in the United States
B. underutilization of relatively small but rich plots of land
C. overexpansion of the world transportation network for shipping agricultural products
D. extension of agrarian depressions beyond national boundaries
E. emphasis on the importance of market forces in determining the process of agricultural products

15. The author implies that the change in the state of the American framers morale during the latter part of the nineteenth century was traceable to the American farmer's increasing perception that the

A. costs of cultivating the land were prohibitive within the United States
B. development of the first transcontinental railroad in the United States occurred at the expense of the American farmer
C. American farming system was about to run out of the new farmland that was required for its expansion
D. prices of American agricultural products were deteriorating especially rapidly on domestic markets
E. proceeds from the sales of American agricultural products on foreign markets were unsatisfactory

16. According to the passage, which of the following occurred prior to 1890?
A. Frederick J. Turner's thesis regarding the American frontier became influential
B. The Homestead Act led to an increase in the amount of newly farmed land in the United States
C. The manufacturers of technologically advanced agricultural machinery rapidly increase their marketing efforts
D. Direct lines of communication were constructed between the United States and South America.
E. Technological advances made it fruitful to farm extensively on a mechanized basis.

17. The author's argument implies that, compared to the yearly price changes that actually occurred on foreign agricultural markets during the 1880s, American farmers would have most preferred yearly price changes that were
A. much smaller and in the same direction
B. much smaller but in the opposite direction
C. slightly smaller and in the same direction
D. similar in size but in the opposite direction
E. slightly greater and in the same direction

APPENDIX C
SPIELBERGER'S STATE ANXIETY INVENTORY

Read each statement and then bubble in the appropriate response to how YOU FEEL RIGHT NOW, THAT IS, AT THIS MOMENT. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer that seems to describe your present feelings best.

1	2	3	4
Not at all	Somewhat	Moderately so	Very much so

1. I feel calm
2. I feel secure.
3. I am tense.
4. I am regretful.
5. I feel at ease.
6. I feel upset.
7. I am presently worrying over possible misfortunes.
8. I feel rested.
9. I feel anxious.
10. I feel comfortable.
11. I feel self-confident.
12. I feel nervous.
13. I am jittery.
14. I feel "high strung."
15. I am relaxed.
16. I feel content.
17. I am worried.
18. I feel over-excited and rattled.
19. I feel joyful.
20. I feel pleasant.

APPENDIX D
STUDY 2 PATRICIPANT SELECTION CRITERIA

Please indicate your age _____ & year in school _____.

Please indicate your typical attitude toward verbally-oriented classes (literature, languages, writing).

Strongly Dislike 0 1 2 3 4 5 6 Strongly Like

If you took the SAT's, please indicate your approximate verbal score.

0. <450
1. 451-500
2. 501-550
3. 551-600
4. 601-650
5. 651-700
6. 701+
7. Did not take the SATs.
8. Cannot remember my score.

APPENDIX E
ANTIIPATED SCORE FOR SELF & ADMINISTRATOR

Participant # ____

You are about to take an test that consists of challenging items extracted from a nationally standardized test of scholastic ability. According to a 1998 nationwide analysis, the average college student responded correctly to approximately 50% of the items. Your responses to the following two items will be kept completely confidential by your sealing this page in an envelope and dropping it into the box by the door on your way out.

- Please indicate the % of items that you anticipate responding to correctly. _____%

- Second, we have asked the experimenter to try to guess how each participant will perform on the test. Try to guess what % of items that the experimenter believes you will score correctly.

_____%

APPENDIX F
STUDY 2 DIFFICULT TEST

SECTION 1
20 Questions, Time - 30 minutes

Directions: Each of the following questions consists of an incomplete sentence followed by five words or pairs of words. Choose that word or pair of words which, when substituted for the blank space or spaces, *best* completes the meaning of the sentence and mark the letter of your choice on your answer sheet.

1. The new biological psychiatry does not deny the contributing roles of psychological factors in mental illness, but that these factors may act as a catalyst on existing physiological conditions and _____ such illnesses.

   (A) disguise
   (B) impede
   (C) constrain
   (D) precipitate
   (E) consummate

2. During periods of social and cultural stability, many art academies are so firmly controlled by _____ that all real creative work must be done by the _____.

   (A) dogmatists . . disenfranchised
   (B) managers . . reactionaries
   (C) reformers . . dissatisfied
   (D) impostors . . academicians
   (E) specialists . .elite

3. The critics were distressed that a young poet of such glowing _____ could descend to writing jingles for TV commercials.

   (A) obscurity
   (B) ill-repute
   (C) attractivenesspromise
   (D) flexibility

4. The first World War began in a context of jargon and verbal delicacy and continued in a cloud of _____ as _____ as language and literature, skillfully used, could make it.

   (A) circumlocution . . literal
   (B) cliché . . lucid
   (C) euphemism . . impenetrable
   (D) particularity . . deliberate
   (E) subjectivity . . enthralling

5. While Parker is very outspoken on issues she cares about, she is not _____; she concedes the _____ of opposing arguments when they expose weaknesses inherent in her own.

   (A) fickle . . validity
   (B) arrogant . . restraint
   (C) fanatical . . strength
   (D) congenial . . incompatibility
   (E) unyielding . . speciousness

Directions: Each of the following questions consists of a pair of words followed by five pairs of words lettered A to E. The words bear some meaningful relationship to each other. Choose the lettered pair of words whose relationship is most similar to that expressed by the original pair and mark its letter below.

6. didactic : instruct ::

   (A) pedantic : contend
   (B) comic : amuse
   (C) theatrical : applaud
   (D) imperative : obey
   (E) rhetorical : recite

7. garrulous : : talkative ::

   (A) suspicious : unreliable
   (B) cantankerous : obtuse
   (C) cloying : sweet
   (D) reflective : insightful
   (E) prudent : indecisive

8. digressive : statement ::

   (A) connotative : definition
   (B) slanderous : slur
   (C) tangential : presupposition
   (D) biased : opinion
   (E) circuitous : route

9. chicanery : clever ::

   (A) expertise : knowledgeable
   (B) certainty : doubtful
   (C) gullibility : skeptical
   (D) machination : heedless
   (E) tactlessness : truthful

10. pluck : quit ::

    (A) verve : flinch
    (B) gall : skimp
    (C) pride : grovel
    (D) charm : smile
    (E) poise : waver

Directions: The passage below is followed by a set of questions. Read the passage, then answer the accompanying questions, basing your answers on what is stated or implied in the passage.

The historian Frederick J. Turner wrote in the 1890's that the agrarian discontent that had been developing steadily in the United States since about 1870 had been precipitated by the closing of the internal frontier—that is, the depletion of available new land needed for further expansion of the American farming system. Not only was Turner's thesis influential at the time, it was later adopted and elaborated by other scholars, such as John D. Hicks in *The Populist Revolt* (1931). Actually, however, new lands were taken up for farming in the United States throughout and beyond the nineteenth century. In the 1890s, when agrarian discontent had become most acute, 1,100,000 new farms were settled, which was 500,000 more than had been settled during the previous decade. After 1890, under the terms of the Homestead Act and its successors, more new land was taken up for farming than had been taken up for this purpose in the United States up until that time. It is true that a high proportion of the newly farmed land was suitable only for grazing and dry farming, but agricultural practices had become sufficiently advanced to make it possible to increase the profitability of farming by utilizing even these relatively barren lands.

The emphasis given by both scholars and statesmen to the presumed disappearance of the American frontier helped to obscure the great importance of changes in the conditions and consequences of international trade that occurred during the second half of the nineteenth century. In 1869 the Suez Canal was opened and the first transcontinental railroad in the United States was completed. An extensive network of telegraph and telephone communications was spun: Europe was connected by submarine cable with the United States in 1866 and with South America in 1874. By about 1870 improvements in agricultural technology made possible the full exploitation of areas that were most suitable for extensive farming on a mechanized basis. Huge tracts of land were being settled and famed in Argentina, Australia, Canada, and in the American West, and these areas were joined with one another and with the countries of Europe into an interdependent market system. As a consequence, agrarian depressions no longer were local or national in scope, and they struck several nations whose internal frontiers had not vanished or were not about to vanish. Between the early 1870s and the 1890s, the mounting agrarian discontent in America paralleled the almost uninterrupted decline in the prices of American agricultural products on foreign markets. Those staple-growing farmers in the United States who exhibited the greatest discontent were those who had become most dependent on foreign markets for the sale of their products. Insofar as Americans had been deterred from taking up new land for farming, it was because market conditions had made this period a perilous time in which to do so.

11. According to the author, changes in the conditions of international trade resulted in

(A) an underestimation of the amount of new land that was being farmed in the United States.
(B) an underutilization of relatively small but rich plots of land.
(C) an overexpansion of the world transportation network for shipping agricultural products.
(D) extension of agrarian depressions beyond national boundaries.
(E) an emphasis on the importance of market forces in determining the process of agricultural products.

12. The author implies that the change in the state of the American framers morale during the latter part of the nineteenth century was traceable to the American farmer's increasing perception that the

(A) costs of cultivating the land were prohibitive within the United States
(B) development of the first transcontinental railroad in the United States occurred at the expense of the American farmer
(C) American farming system was about to run out of the new farmland that was required for its expansion
(D) prices of American agricultural products were deteriorating especially rapidly on domestic markets
(E) proceeds from the sales of American agricultural products on foreign markets were unsatisfactory

13. According to the passage, which of the following occurred prior to 1890?

(A) Frederick J. Turner's thesis regarding the American frontier became influential.
(B) The Homestead Act led to an increase in the amount of newly farmed land in the United States.
(C) The manufacturers of technologically advanced agricultural machinery rapidly increase their marketing efforts.
(D) Direct lines of communication were constructed between the United States and South America.
(E) Technological advances made it fruitful to farm extensively on a mechanized basis.

14. The author's argument implies that, compared to the yearly price changes that actually occurred on foreign agricultural markets during the 1880s, American farmers would have most preferred yearly price changes that were

(A) much smaller and in the same direction
(B) much smaller but in the opposite direction
(C) slightly smaller and in the same direction
(D) similar in size but in the opposite direction
(E) slightly greater and in the same direction

Directions: Each of the following questions consists of an incomplete sentence followed by five words or pairs of words. Choose that word or pair of words which, when substituted for the blank space or spaces, *best* completes the meaning of the sentence and mark the letter of your choice on your answer sheet.

15. Despite her _____ unwillingness, the promoters were still hopeful that, given sufficient diplomacy and flattery on their part, they could _____ her into signing the recording contract.

    (A) patent . . entrap
    (B) extreme . . intimidate
    (C) apparent . . shame
    (D) painful . . tantalize
    (E) obvious . . inveigle

16. Usually the first to spot data that were inconsistent with other findings, in this particular experiment she let a number of _____ results slip by.

    (A) inaccurate
    (B) verifiable
    (C) redundant
    (D) salient
    (E) anomalous

17. An experienced politician who knew better than to launch a campaign in troubled political waters, she intended to wait for a more _____ occasion.

    (A) provocative
    (B) unseemly
    (C) questionable
    (D) advantageous
    (E) theoretical

18. Cèzanne's delicate watercolor sketches often served as _____ of a subject, a way of gathering fuller knowledge before the artist's final engagement of the subject in an oil painting.

    (A) an abstraction
    (B) an enchantment
    (C) a synthesis
    (D) a reconnaissance
    (E) a transcription

19. Liberty is not easy, but far better to be an _____ fox, hungry and threatened on its hill, than a _____ canary, safe and secure in its cage.

    (A) unfriendly . . fragile
    (B) aging . . young
    (C) angry . . content
    (D) imperious . . lethargic
    (E) unfettered . . well-fed

20. Johnson never _____ to ignore the standards of decent conduct mandated by company policy if _____ compliance with instructions from his superiors enabled him to do so, whatever the effects on his subordinates.

    (A) deigned . . tacit
    (B) attempted . . half-hearted
    (C) intended . . direct
    (D) scrupled . . literal
    (E) wished . . feigned

STOP
END OF SECTION 1

APPENDIX G
STUDY 2 EASY TEST

SECTION 1
20 Questions, Time - 30 minutes

Directions: Each of the following questions consists of an incomplete sentence followed by five words or pairs of words. Choose that word or pair of words which, when substituted for the blank space or spaces, *best* completes the meaning of the sentence and mark the letter of your choice on your answer sheet.

1. The expansion of the military strength of our nearby enemy fills us with _____.

   (A) guilt
   (B) amazement
   (C) admiration
   (D) apprehension
   (E) disgust

2. While a great deal of change and modernization has taken place in India since 1947, the basic economic arrangements, values, and family roles have been generally _____.

   (A) overturned
   (B) stable
   (C) modified
   (D) complicated
   (E) appropriate

3. The critics were distressed that a young poet of such glowing _____ could descend to writing jingles for TV commercials.

   (A) obscurity
   (B) ill-repute
   (C) attractiveness
   (D) promise
   (E) flexibility

4. Joy kept here late parents' furniture, not for any _____ value it had, but for purely _____ reasons.

   (A) potential...monetary
   (B) ornamental...aesthetic
   (C) financial...commercial
   (D) inherent...sentimental
   (E) personal...accidental

5. It is the task of the International Wildlife Preservation Commission to prevent endangered species from becoming _____ in order that future generations may _____ the great diversity of animal life.

   (A) tamed...recollect
   (B) evolved...value
   (C) extinct...enjoy
   (D) specialized...anticipate
   (E) widespread...appreciate

6. It is remarkable that a man so in the public eye, so highly praised and imitated, can retain his _____.

   (A) pretentiousness
   (B) dogmas
   (C) bravado
   (D) idiosyncrasies
   (E) humilit

7. I regret that we shall be unable to _____ all who have requested rooms for the weekend of graduation.

   (A) correlate
   (B) accommodate
   (C) explicate
   (D) saturate
   (E) intimidate

8. For all the _____ involved in the study of seals, the Arctic researchers have occasional moments of pure _____ over some new idea or discovery.

   (A) tribulations...despair
   (B) hardships...exhilaration
   (C) confusions...bewilderment
   (D) inconvenience...panic
   (E) thrills...delight

9. Even critics who do not _____ Robin Williams' interpretation of the part in his last movie _____ him as an inventive comic actor.

   (A) dissent from...dismiss
   (B) cavil at...welcome
   (C) concur with...denounce
   (D) recoil from...deride
   (E) agree with...acknowledge

10. The credit manager said the company would report to private credit bureaus and collection agencies the names of more than two thousand _____ borrowers, whom he described as "deadbeats."

    (A) potential
    (B) delinquent
    (C) acceptable
    (D) insubstantial
    (E) esteemed

11. Those interested in learning the latest findings about how drug use affects the brain will have to _____ the excellent technical journals where most of the important findings are _____.

    (A) subscribe to...ignored
    (B) suffer through...located
    (C) rely on...unrepresented
    (D) resort to...found
    (E) complain about...published

12. Conservative Republicans argue that restrictions on gun ownership will _____ rather than alleviate crime.

    (A) resolve
    (B) diminish
    (C) condemn
    (D) minimize
    (E) exacerbate

13. Archeologists are involved in _____ Mayan temples in Central America, uncovering the old ruins in order to learn more about the civilization they represent.

    (A) demolishing
    (B) incapacitating
    (C) excavating
    (D) worshipping
    (E) adapting

14. Afraid that the _____ nature of the plays being presented would corrupt the morals of their audiences, the Puritans closed the theaters in 1642.

    (A) mediocre
    (B) fantastic
    (C) profound
    (D) lewd
    (E) witty

15. The incidence of smoking among women, formerly _____, has grown to such a degree that lung cancer, once a minor problem, has become the chief _____ of cancer-related deaths among women.

(A) negligible...cancer
(B) minor...antidote
(C) pre-eminent...cure
(D) relevant...modifier
(E) pervasive...opponent

16. To the relief of the archaeologists, although two of the three sacred burial sites had been _____ by animals rummaging through them, one was _____.

(A) destroyed... spoiled
(B) disturbed...intact
(C) verified...existent
(D) impaired... fragmented
(E) elevated...desecrated

17. An experienced politician who knew better than to launch a campaign in troubled political waters, she intended to wait for a more _____ occasion.

(A) provocative
(B) unseemly
(C) questionable
(D) advantageous
(E) theoretical

18. Quick-breeding and immune to most pesticides, cockroaches are so _____ that even a professional exterminator may fail to _____ them.

(A) vulnerable... remove
(B) widespread...notice
(C) fragile...destroy
(D) hardy...eliminate
(E) numerous...detect

19. Joe bore the pain _____ and did not wince or whimper when the incision was made.

(A) hysterically
(B) stoically
(C) poorly
(D) marginally
(E) gleefully

20. The number of black hawks has _____ because the intrusion of humans on their territory have caused them to _____ their customary breeding places.

(A) multiplied...endure
(B) enhanced...alter
(C) increased...locate
(D) diminished...accept
(E) dwindled...shun

APPENDIX H
COGNITIVE INTERFERENCE QUESTIONNAIRE

Directions: Before you begin SECTION 2 of the test, please respond to the following items as honestly as possible. Your responses will be completely anonymous and any information provided will be held in strict confidentiality.

Please indicate below how often the following thoughts occurred to you during the test.

1	2	3	4	5
I had this thought very often				I never had this thought

____ 1. I thought about how poorly I was doing.
____ 2. I wondered what the experimenter would think of me.
____ 3. I thought about how I should be working more carefully.
____ 4. I thought about how much time I had left.
____ 5. If I don't do well on this test, others may question my ability.
____ 6. I thought about how others have done on the test.
____ 7. I thought about things that were not useful to solving the test questions.
____ 8. If I do poorly on this test, people will look down on me.
____ 9. I thought about the difficulty of the problems.
____ 10. I thought about my level of ability.
____ 11. I thought about the purpose of the experiment.
____ 12. I thought about how I would feel if I were told how I performed on the test.
____ 13. I thought about how often I got confused.
____ 14. I thought about things completely unrelated to the experiment.
____ 15. I thought about whether the test was unfair.
____ 16. I thought about what the other participants were thinking of me.
____ 17. I thought that the experimenter does not believe that I will do well on the test.
____ 18. I felt that I was doing well on the test.
____ 19. I thought that I may not have the ability to do well on this test.
____ 20. People will think I have less ability than I do if I do not do well on this test.

APPENDIX I
REVISED FEAR OF NEGATIVE EVALUATION SCALE

Please indicate below the extent to which the following descriptions are characteristic of you.

1	2	3	4	5
Extremely Uncharacteristic of Me				Extremely Characteristic of Me

____ 21. I worry about what people will think of me even when I know it doesn't make any difference.

____ 22. I am unconcerned even if I know people are forming an unfavorable impression of me.

____ 23. I am frequently afraid of other people noting my shortcomings.

____ 24. I rarely worry about what kind of impression I am making on someone.

____ 25. I am afraid that others will not approve of me.

____ 26. I am afraid that people will find fault with me.

____ 27. When I am talking to someone, I worry about what they may be thinking about me.

____ 28. I am usually worried about what kind of impression I make.

____ 29. If I know someone is judging me, it has little effect on me.

____ 30. Sometimes I think I am too concerned with what other people think of me.

____ 31. I often worry that I will say or do the wrong things.

____ 32. I'm concerned about the way I present myself.

____ 33. I'm concerned about my style of doing things.

APPENDIX J
SELF-CONSCIOUSNESS SCALE

Please indicate below the extent to which the following descriptions are characteristic of you.

1	2	3	4	5
Extremely Uncharacteristic of Me				Extremely Characteristic of Me

____ 34. I'm concerned about the way I present myself.

____ 35. I'm self-conscious about the way I look.

____ 36. I usually worry about making a good impression.

____ 37. One of the last things I do before I leave my house is look in the mirror.

____ 38. I'm concerned about what other people think of me.

____ 39. I'm usually aware of my appearance.

____ 40. I'm always trying to figure myself out.

____ 41. Generally, I'm not very aware of myself.

____ 42. I reflect about myself a lot.

____ 43. I'm often the subject of my own fantasies.

____ 44. I never scrutinize myself.

____ 45. I'm generally attentive to my inner feelings.

____ 46. I'm constantly examining my motives.

____ 47. I sometimes have the feeling that I'm off somewhere watching myself.

____ 48. I'm alert to changes in my mood.

____ 49. I'm aware of the way my mind works when I work through a problem.

APPENDIX K
SHORT SPIELBERGER STATE ANXIETY INVENTORY

Read each statement and then bubble in the appropriate response to how YOU FEEL RIGHT NOW, THAT IS, AT THIS MOMENT. There are no right or wrong answers. Don not spend too much time on any one statement but give the answer that seems to describe your present feelings best.

1	2	3	4
Not at all	Somewhat	Moderately so	Very much so

____ 50. I am tense

____ 51. I am regretful

____ 52. I feel upset

____ 53. I am presently worrying over possible misfortunes

____ 54. I feel anxious

____ 55. I feel nervous

____ 56. I am worried

____ 57. I feel rattled.

APPENDIX L
POSITIVE AND NEGATIVE AFFECT SCALE

The following scale consists of a number of words that describe different feelings and emotions. Read each item and indicate to what extent you feel this way right now. Use the following scale to record your answers.

1	2	3	4	5
Very slightly or None at all	A little	Moderately	Quite a bit	Extremely

____ interested	____ distressed
____ excited	____ upset
____ strong	____ scared
____ hostile	____ enthusiastic
____ proud	____ distracted
____ ashamed	____ inspired
____ determined	____ attentive

APPENDIX M
OPEN-ENDED THOUGHT-LISTINGS

Please indicate below in as much detail as you'd like any thoughts you are having now or that were going through your mind during the first half of the test. Also include any thoughts you may be having about your performance on the test, the test itself, or more general aspects of the study such as the experimenter or the setting.

APPENDIX N
EXPERIMENTER TRAIT QUESTIONNAIRE

Dear Participant,

The primary objective of the Institutional Review Committee at the University of Florida is to maintain the quality of the experimental process.

While you have most likely completed this evaluation for previous studies, we are interested in your honest evaluation of the research assistant who conducted the experiment that you have just completed. We know that the participant's perceptions of the experimenter can have an important impact on the outcome of the study. Your responses will remain confidential so we ask that you express your honest opinion.

1	2	3	4	5	6	7	8
Strongly Disagree							Strongly Agree

1. The experimenter communicated instructions for the study to you in a clear and concise manner. _____

2. The experimenter showed appropriate respect and concern for you. _____

3. The experimenter was inattentive to you. _____

4. The experimenter appeared to be shy and reserved. _____

5. The experimenter seemed arrogant. _____

6. The experimenter seemed to be the type of person who might evaluate person unfairly based on race, gender or sexual preference. _____

7. The University of Florida is considering instituting a research mentorship program between graduate students and undergraduate students. To what extent would you like to be matched with this experimenter as your mentor on future endeavors? _____

Any comments.

_____

_____

REFERENCES

Aiken, L., & West, S. (1993). Multiple regression: Testing and interpreting interactions. Newbury Park, CA: Sage.

Allport, G. (1954). The nature of prejudice. New York: Addison-Wesley.

Aronson, J., Lustina, M., Good, C., Keough, K., Steele, C., & Brown, J. (1999). When White men can't do math: Necessary and sufficient conditions for stereotype threat. Journal of Experimental Social Psychology, 35, 29-46.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. Psychological Review, 84, 191-215.

Baron, S. (1986).Distraction/conflict theory: Progress and problems. In L. Berkowitz (Ed.), Advances in experimental social psychology, Vol.19. New York: Academic Press.

Baumeister, R., & Showers, C. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. European Journal of Social Psychology, 16, 361-383.

Blascovich, J., Quinn, D., Spencer, S., & Steele, C. (1998). The effect of stereotype threat on African Americans' blood pressure. Unpublished manuscript, University of California, Santa Barbara.

Bond, C. (1982). Social facilitation: A self-presentational view. Journal of Personality and Social Psychology, 42, 1042-1050.

Brown, L. (1998). Ethnic stigma as a contextual experience: A possible selves perspective. Personality and Social Psychology Bulletin, 24, 163-172.

Brown, L., Boniecki, K., & Walters, A. (1998). The group schema versus the group stereotype: Tolerant and prejudiced views of African Americans. Unpublished manuscript, University of Florida, Gainesville.

Cheryan, S., & Bodenhausen, G. (2000). When positive stereotypes threaten intellectual performance: The psychological hazards of "Model Minority" status. Psychological Science, 11, 399-402

Clark, K., & Clark, M. (1958). Racial identification and preference in Negro children. In E. E. Maccoby, T. M. Newcomb, and E. E. Hartley (Eds.), Readings in social psychology, 3rd Edition. New York: Holt and Company.

Cottrell, N., Wack, D., Sekerak, G., & Rittle, R. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. Journal of Personality and Social Psychology, 9, 245-250.

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. Psychological Review, 96, 608-630.

Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. Journal of Personality and Social Psychology, 60, 218-228.

Croizet, J., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. Personality and Social Psychology Bulletin, 24, 588-594.

Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. Journal of Personality and Social Psychology, 56, 5-18.

Fenigstein, A, Scheier, M., & Buss, A. (1975). Public and private self-consciousness: Assessment and theory. Journal of Consulting and Clinical Psychology, 43, 522-527.

Geen, R. (1989). Alternative conceptions of social facilitation. In P. Paulus (Ed.), Psychology of group influence, (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Goffman, E. (1959). The presentation of self in everyday life. Garden City, NY: Doubleday.

Goffman, E. (1963). Stigma: Notes on the management of spoiled identity. New York: Simon & Shuster, Inc.

Herrnstein, R., & Murray, C. (1994). The bell curve. New York: Grove Press.

Howell, D. C. (1982). Statistical methods for psychology. Duxbury Press: Boston.

Hull, J., Tedlie, J., & Lehn, D. (1992). Moderator variables in personality research: The problem of controlling for plausible alternatives. Personality and Social Psychological Bulletin, 18, 115-117.

Inzlicht, M, & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem solving deficits in the presence of males. Psychological Science, 11, 365-370.

Jencks, C., & Phillips, M. (1998). The Black-White test score gap. Cambridge, MA.

Kleck, R., & Strenta, A. (1980). Perceptions of the impact of negatively physical characteristics on social interaction. Journal of Personality and Social Psychology, 39, 861-873.

Leary, M. (1983). A brief version of the Fear of Negative Evaluation Scale. Personality and Social Psychology Bulletin, 9, 371-375.

Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. Journal of Personality and Social Psychology, 71, 1092-1107.

Lewin. K. (1948). Resolving social conflicts. New York: Harper and Brothers Publishers.

Major, B., & Schmader, T. (1998). Coping with stigma through psychological disengagement. In J. Swim & C. Stangor (Eds.), Stigma: The target's perspective. New York: Academic Press.

Miller, L. (1995). An American imperative: Accelerating minority educational advancement. New Haven, CT: Yale University Press.

Niemann, Y. F., Jennings, L., Rozelle, R. M., & Baxter, J. C. (1994). Use of free responses and cluster analysis to determine the stereotypes of eight groups. Personality & Social Psychology Bulletin, 20, 379-390.

Neisser, U., Boodoo, , G., Bouchard, T. J., Boykin, W. A., Brody, N., Ceci, S. J., Halpern, D. F., Loehin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. American Psychologist, 51, 77-101.

Ogbu, J. (1986). The consequences of the American caste system. In U. Neisser (Ed.), The school achievement of minority children: New perspectives. Hillsdale, NJ: Erlbaum.

Pettigrew, T. (1964). A profile of the Negro American. Princeton, NJ: D. Van Nostrand Company, Inc.

Pilotte, W., Gable, R. (1990). The impact of positive and negative anxiety stems on the validity of a computer anxiety scale. Educational & Psychological Measurement, 50, 603-610.

Pyszczynski, T., & Greenberg, J. (1983). Determinants of reduction in effort as a strategy for coping with anticipated failure. Journal of Research in Personality, 17, 412-422.

Sapp, M., Durand, H., & Farrell, W. (1995). Measures of actual test anxiety in educationally and economically disadvantaged college students. College Student Journal, 29, 65-72.

Sarason, I. (1972). Experimental approaches to test anxiety: Attention and the uses of information. In C. D. Spielberger (Ed.), Anxiety: Current trends in theory and research, Vol.2. New York: Academic Press.

Sarason, I. (1980). Introduction to the study of test anxiety. In I. G. Sarason (Ed), Test anxiety: Theory, research, and applications (pp. 3-13). Hillsdale, NJ: Erlbaum.

Schlenker, B. (1980). Impression management: The self-concept, social identity, and interpersonal relations. Monterey, CA: Brooks/Cole.

Schlenker, B., & Leary, M. (1982). Social anxiety and self-presentation: A conceptualization and model. Psychological Bulletin, 92, 641-669.

Schotte, C., Maes, M., Cluydts, R., & Closyns, P. (1996). Effects of affective-semantic mode of item presentation in balanced self-report scales: Biased construct validity of the Zung Self-rating Depression Scale. Psychological Medicine, 26, 1161-1168

Spencer, S., Steele, C., & Quinn, D. (1999). Stereotype threat and women's math performance. Journal of Experimental Social Psychology, 35, 4-28.

Spielberger, C., Gorsuch, R., & Lushene, R. (1970). Manual of the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologist Press.

Steele, C. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. American Psychologist, 52, 613-629.

Steele, C., & Aronson, J. (1995). Stereotype threat and the intellectual performance of African Americans. Journal of Personality and Social Psychology, 69, 797-811.

Stone, J., Lynch, C., Sjomeling, M., & Darley, J. (1999). Stereotype threat effects on Black and White athletic performance. Journal of Personality and Social Psychology, 77, 1213-1227.

Watson, D., Clarke, L., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scale. Journal of Personality and Social Psychology, 57, 672-680.

Wine, J. (1971). Test anxiety and direction of attention. Psychological Bulletin, 76, 92-104.

Zajonc, R. (1965). Social facilitation. Science, 149, 269-274.

BIOGRAPHICAL SKETCH

Alyssa Walters received her Bachelor of Arts in Psychology from the University of Texas at Austin in 1992. In 1993, Alyssa received a Master of Education in Counseling and Guidance from Southwest Texas State University in San Marcos, Texas. She worked as a Research Associate at the Uniformed Services University of the Health Sciences in Bethesda, Maryland for the two years prior to beginning her graduate studies at the University of Florida. At the University of Florida, Alyssa earned a Master of Social Psychology in 1998, and she will earn a Doctor of Philosophy in 2000.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James A. Shepperd, Chairman
Associate Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Lisa M. Brown
Assistant Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Martin Heesacker
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mary Howard-Hamilton
Associate Professor of Educational Leadership, Policy, and Foundations

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Barry R. Schlenker
Professor of Psychology

This dissertation was submitted to the Graduate Faculty of the Department of Psychology in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2000

_____

Dean, Graduate School